REPUBLIQUE ALGERIENNE DEMOCRATIQUE & POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR & DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE DJILLALI LIABES
FACULTE DES SCIENCES
SIDI BEL-ABBÈS

BP 89 SBA 22000 –ALGERIE- TEL/FAX 048-54-43-44

THESE

Présentée par:
HEBCHI CHAIMA

Pour obtenir le Diplôme de Doctorat

Spécialité : Mathématiques
Option : Probabilités-Statistiques

Intitulée

Propriétés asymptotiques uniforme de l'estimation non paramétrique du mode régression par la méthode locale linéaire fonctionnelle

Thèse soutenue le ***13 December 2020***
Devant le jury composé de :

Président :
Mr BENAISSA Samir Professeur à L'Université S.B.A

Directeur de thèse :
Mr CHOUAF Abdelhak Maitre de Conférence A à L'Université S.B.A.
Examinateur :
Mr MECHAB Boubaker Professeur à L'Université S.B.A
Examinateur :
Mr AZZOUZI Badreddine Maître de Conférence A à E.S.M de Tlemcen

..........To my parents and my family members.

............To my sweetheart HAMDI Malak.

Acknowledgements

I am deeply grateful to my advisor Dr CHOUAF Abdelhak for the continuous support of my Ph.D study and research, for his patience, motivation,for his continuous encouragement. I really appreciate his comments and suggestions which have been extremely valuable for the completion of this task.

Besides my advisor, I would like to offer my special thanks to the examination committee members : Prof. BENAISSA Samir, Prof. MECHAB Boubaker and Dr. AZZOUZI Badreddine for having devoted some of their precious time to very careful readings of this thesis.

Contents

Chapter 1

Introduction

It is universally acknowledged that technology plays a vital role in the human life and in many domains especially in science, so the development of technology gives us many studies in different fields by modern and relevant measuring instruments. In general, the observations of data in the majority of applied sciences have a functional properties (the data are surfaces or curves for instance : electricity consumption data, the spectrometric data, ...) therefore, the functional data analysis appears to model and treat such kind of data, for an updated of references, we can lead the reader to the monographs by Ramsay and silverman (2002, 2005), Bosq (2000) and Ferraty and Vieu (2006).

Conditional mode, conditional quantiles and conditional median are the most popular tools which have the feature of summarizing data. These last predictors have a strong relationship with conditional distribution and its derivatives. In nonprametric estimating and for functional data, Ferraty et al.(2006) introduced the estimator model of conditional distribution and its derivatives, the authors established the rate of convergence (almost complete convergence). This study seems like the starting point of many studies (eg., Lacksaci et al.(2008), Ezzahrioui and Ould-Saïd (2008), Lacksaci et al.(2012),...).

For more than decades, many papers relied on the kernel method to estimate the nonparametric regression function (see : Rosenblatt (1969), Ramsay and Dalzell (1991) and Hastie and Mallows (1993)), then Ferraty and Vieu (2000) generalized the kernel regression estimator of Nadaraya-Watson familiar function where this model was adopted in many studies to find more asymptotic results such as : the k-nearest-neighbours (k-NN) estimator is investigated by Burba et al.(2009), convergence in $\mathbb{L}^2$ norm (see : Dabo-Niang and Rhomani (2003)) which is generalized in the α-mixing case (see Delsol (2007)) and recently, Kara-Zaitri et al. (2017) stated the uniform in bandwidth for kernel regression estimator.

When the derivatives estimator of regression come down to providing us about the behaviour of both regression shape and regression mode, Mack and Müller (1989) used the Nadaraya-Watson kernel estimator of the v-th ($v \geq 0$) derivative of a regression function to establish some consistency results for instance : the asymptotic normality, the asymptotic mean squared error (AMSE) also the choice of bandwidth was discussed. M convergence of regression function estimator and its derivatives was studied by Boularan et al. (1995) and we point out that the reader can find important results in ρ-mixing case and α-mixing case by taking the monograph by Györfi et al. (1989). In mode regression aspect, Lee (1993) introduced the mode regression with quadratic kernel (QME)in the case of a truncated dependent variable and Ziegler (2002) established the probability convergence and the normality asymptotic of mode regression, this study is based on Nadaraya-Watson kernel estimator for scalar explanatory variable X and for the higher derivatives of regression estimator. For recent studies, we can see : Kemp and Santos Silva (2012) and Chen et al. (2017).

However, the previous literatures used the Nadaraya-Watson techniques as estimation method which has some drawbacks, mainly, in the bias term. Hence, in the functional data setup, the local linear method comes to generalize and ameliorate the kernel method. Actually, Baíllo and Grané (2009) proposed the first local linear estimator model of the regression operator when the explanatory variable takes values in a Hilbert space. When the regressors take values in semi-metric space, Barrientos-Marin et al. (2010) introduced another version of the local linear estimator of the regression operator. This last method has been extended to estimate the conditional distribution and its derivatives (Demongeot el al.(2013), Rachdi et al. (2014), Demongeot el al.(2014) and Messaci et al. (2015)).

This thesis which sheds more light on the uniform asymptotic behaviours of both the local linear estimation of conditional cumulative distribution function and the local linear mode regression for functional data (without forgetting the kernel estimation of mode regression), has been organized in the following chapters :

- Chapter 2 "Statistical background and general notations for nonparametric functional data"
  In this chapter, we will deal on the statistical background and general notations for nonparametric functional data so, we attempt to investigate different aspects which are related to join the nonparametric statistics with functional data, as well as, the definition of nonparametric statistics for functional data also we provide a brief definition, aims of functional data analysis (FDA) in addition, the distinction between functional variable and functional datasets. Moreover, to make the previous vocabulary more clear, we selected two statistical problems (spectrometric data and phonetic data) for more explanation and clarification we refer to the monograph of Ramsay and Silverman (2002) and Ferraty and Vieu (2006). As the asymptotic properties need some tools and notations, we took the semi-metric as better space, small ball probabilities as important measure in the different hypotheses on rate of convergence (see Ferraty and Vieu (2006)). Furthermore, we use the complete convergence notion as a tool to obtain our asymptotic results and Bernstein's inequality as the easiest inequality. At the end of chapter, we present some operator models and provide them by their estimators using local linear method and kernel method.

- Chapter 3 "Some asymptotic results linked with nonparametric estimation for functional covariates "
  What we are mainly concerned here in this chapter is the study of some results of many papers and gives some asymptotic results. Based on the same notations and models of previous chapter and by using two kinds of estimation methods (the kernel method and the local linear method), the present chapter consists of three sections; the first one tackles the main related information about the regression operator, the first part of this section highlights the importance of the kernel method theory. More accurately, we concentrate on the asymptotic behaviour of kernel regression estimator for instant: the pointewise almost complete convergence, the rate of con-

vergence and also the p-mean convergence. On the other side, the second part gives a brief summary of main results linked with the local linear regression estimator, we adopt the model of Barrientos et al.(2010) (which we rely heavily on the last chapter) to bring the principal findings like : the rate of pointewise almost complete convergence, the mean-squared convergence and the asymptotic normality. For the same kind of consistency results of the first section, we discuss the second section for the conditional cumulative distribution function and the conditional density function for the last section.

- Chapter 4 "Local linear estimation of c.d.f in the functional data: Uniform consistency with convergence rates"
  Through this chapter, we intend to spotlight on the nonparametric estimation of conditional cumulative distribution function when the explanatory variable X has a functional nature and for scalar response Y. We try to extend the results of Demongeot et al.(2014) and bearing in mind the outcome of Demongeot et al.(2013) to get the rate of the uniform pointwise almost complete convergence of our estimator. Hence, we adopt the model and the estimators of Demongeot et al.(2014), also under some technical assumptions and topological structure we state our main theorem which can be deduced from three lemmas and one corollary. Furthermore, we provide our results by studying the conditional mode estimation and of course the appendix section is devoted to the lemmas' proofs.

- Chapter 5 "Kernel estimation of mode regression for functional data"
  In this chapter, we deal with the functional mode regression in order to highlight the relationship between an explanatory functional variable X and a scalar response Y in iid setting, moreover, we attempt to establish the pointwise almost complete convergence and the almost complete convergence rate of mode regression estimator by using kernel method for functional data.

  The structure of this chapter is as follows : in section 1, we present our models and estimators, under some conditions, section 2 is devoted to state the main results. A simulation study carried out in section 3 and proofs can be found in the last section.

- Chapter 6 "Uniform almost complete convergence of local linear mode regression"
  this chapter aims to joint the merits of mode with regression function. So, we study the derivatives of regression operator to identify the behaviour of mode. In i.i.d. setting and for functional data, the present chapter is provided to establish the rate of uniform almost complete convergence of mode regression by using local linear method. In this study, we attempt to extend the results of Barrientos et al.(2010), therefore, we take the same model of earlier authors and try to estimate the first derivative of regression operator. Under ten conditions which are commonly used in many studies of the local linear method for functional data we establish the main result, the theorem's proof can be deduced directly from three lemmas which proved in the appendix section.

- Chapter 7 "Conclusion and perspectives"
  The final chapter is putted forward to sum up some results gained from our investigation and followed by some perspectives in order to improve our research.

References

BAÍLLO, A. and GRANÉ, A. (2009). Functional Local Linear Regression with Functional Predictor and Scalar Response. *Journal of Multivariate Analysis*, **100**, 102-111.

BARRIENTOS-MARIN, J., FERRATY, F. and VIEU, P .(2010). Locally Modelled Regression and Functional Data.*J. of Nonparametric Statistics, **22**, 617-632.*

BOSQ, D. *(2000). Linear processes in function spaces. Theory and applications.* Lecture Notes in Statistics, **149.** Springer-Verlag.

BOULARAN, J., FERRÉ, L., VIEU, P. (1995). Location of particular points in nonparametric regression analysis. *Austral. J. Statist.* **37**, 161-168.

BURBA, F., FERRATY, F. and VIEU, P . (2009). k-Nearest Neighbour method in functional nonparametric regression.*Journal of Nonparametric Statistics, **21**, 453-469.*

CHEN, Y., MA, X. *and* ZHOU, J. *(2017). Variable selection for mode regression.* Journal of Applied Statistics, **45**, 1-8.

DABO-NIANG, S. and RHOMARI, N. (2003). Kernel regression estimation when the regressor takes values in metric space. *C. R. Acad. Sci. Paris*, **336**, 75-80.

DELSOL, L. (2007). CLT and $\mathbb{L}^q$ errors in nonparametric functional regression, *C. R. Math. Acad. Sci, **345** (7), 411-414.*

DEMONGEOT, J., LAKSACI, A., MADANI, F. *and* RACHDI, M. *(2013). Functional data : local linear estimation of the conditional density and its application.*Statistics*, **47**, 26-44.*

DEMONGEOT, J., LAKSACI, A. RACHDI, M. *and* RAHMANI, S. *(2014). On the local Modalization of the conditional distribution for functional data.* Sankhya A*, **76** (2), 328-355.*

EZZAHRIOUI, M. *and* E. OULD SAÏD*. (2008). Asymptotic normality of a non-parametric estimator of the conditional mode function for functional data.*J. of Nonparametric Stat, **20**, 3-18.

EZZAHRIOUI, M. and OULD SAÏD, E. (2010).Some asymptotic results of a non-parametric conditional mode estimator for functional time series data. *Neerlandica, **64**, 171-201.*

FERRATY, F. *and* VIEU, P. *(2000). Dimension fracale et estimation de la régression dans des espaces vectoriels semi-normés,* Compte Rendus de l'Académie des Sciences Paris, **330**, 403-406.

FERRATY, F.,LAKSACI, A. andVIEU, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models. *Statistical Inference for Stochastic Processes*, **9**, 47-76.

FERRATY, F. and VIEU, P. (2006). *Nonparametric functional data analysis : Theory and Practice. Springer Series in Statistics, New York : Springer-Verlag.*

GYÖRFI, L., HÄRDLE, W., SARDA, P. *and* VIEU,P*. (1989).* Nonparametric Curve Estimation from Time Series. Springer-Verlag New York.

HASTIE, T. and MALLOWS, C. (1993). Discussion of "A statistical view of some chemometrics regression tools".*Technometrics, **35**, 140-143.*

KARA, L-Z., LAKSACI, A., RACHDI, M. *and* VIEU, P. *(2017). Uniform in bandwidth consistency for various kernel estimators involving functional data.* Journal of Nonparametric Statistics, **29**, 85-107.

KEMP, G. and SANTOS SILVA, J. (2012). Regression towards the mode.*Journal of Econometrics, **170**, 92-101.*

LAKSACI, A., LEMDANI, M. *and* OULD-SAÏD, E*. (2008). A generalized L^1-approach for a kernel estimator of conditional quantile with functional regressors : consistency and asymptotic normality.*Stat. Probab. Lett, **79**, 1065-1073.

LAKSACI, A., MADANI, F. and RACHDI, M. (2012). Kernel conditional density estimation when the regressor is valued in a semi metric space. *Communications in Statistics-Theory and Methods, **42**(19), 3544-3570.*

LEE, M. *(1993). Quadratic mode regression.*J. Econometrics, **57**, 1-19.

MACK ,Y. P. and MÜLLER H. G. (1989). Derivative Estimation in Nonparametric Regression with Random Predictor Variable.*Sankhyā, **51**(1), 59-72.*

MESSACI, F., NEMOUCHI, N., OUASSOU, I. *and* RACHDI, M. *(2015). Local polynomial modelling of the conditional quantile for functional data.*Statistical Methods and Applications, **24**(4), 597-622.

RACHDI, M., LAKSACI, A.,DEMONGEOT, J., ABDALI, A. and MADANI, F. (2014). Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data. *Computational Statistics and Data Analysis*, **73**, 53-68.

RAMSAY, J. and DALZELL, C. (1991). Some tools for functional data analysis.*J. R. Statist. Soc. B, **53**, 539-572.*

RAMSAY, J. O. *and* SILVERMAN, B.W. *(2002).* Applied functional data analysis: Methods and Case Studies. Springer Series in Statistics. Springer-Verlag, New York.

RAMSAY, J. O. and SILVERMAN, B.W. (2005). *Functional Data Analysis. 2nd ed.Springer, New-York.*

ROSENBLATT, M. *(1969). M. Conditional probability density and regression estimators.*in Multivariate Analysis II, Ed. P.R. Krishnaiah,, **25-31**, New York: Academic Press.

ZIEGLER, K. (2002). On nonparametric kernel estimation of the mode of the regression function in the random design model. *Journal of Nonparametric Statistics*, **14**(6), 749-774.

Chapter 2

Statistical background and general notations for nonparametric functional data

In recent years, there has been an increasing interest in functional nonparametric statistics, this field of research rely mainly on some mathematical vocabulary statistics, basic definitions and notations. So, in the sequel, we are going to fix some notations and tools that are crucial to state the asymptotic results of the remaining chapters. At the end of this chapter, we will introduce some model of operators in both methods : the kernel and local linear.

2.1 Nonparametric statistics for functional data

There is general agreement that the ability of technolology developpement to collect and store data leads the staticians to look at the necessity to improve statistical methods or models (models that are suited for curves) in order to take into account the functional structure of this kind of data.

The appellation of functional nonparametric statistics refers to the form and the nature of the set of constraints and the data respectively.

In short, nonparametric derives frome the infinite dimensional feature of the object to be estimated and functional designation comes from the infinite dimensional feature of the data. For more clearly, we give the following definitions

Definition 2.1 *(Ferraty and Vieu (2006))*
Let $\boldsymbol{X}$ be a random vector valued in $\mathbb{R}^p$ and let φ be a function defined on $\mathbb{R}^p$ and depending on the distribution of $\boldsymbol{X}$. A model for the estimation of φ consists in introducing some constraint of the form

$$\varphi \in \mathcal{C}.$$

The model is called a parametric model for the estimation of φ if $\mathcal{C}$ is indexed by a finite number of elements of $\mathbb{R}$. Otherwise, the model is called a nonparametric model.

It is clear that the previous definition makes the difference between parametric and nonparametric models, and by extending this definition to the functional framework, we can get the following definition :

Definition 2.2 *(Ferraty and Vieu (2006))*
Let Z be a random variable valued in some infinite dimensional space $\mathcal{F}$ and let φ be a mapping defined on $\mathcal{F}$ and depending on the distribution of Z. A model for the estimation of φ consists in introducing some constraint of the form

$$\varphi \in \mathcal{C}.$$

The model is called a functional parametric model for the estimation of φ if $\mathcal{C}$ is indexed by a finite number of elements of $\mathcal{F}$. Otherwise, the model is called a functional nonparametric model.

2.2 Functional data analysis (FDA)

Historically, the functional extension of the principal component analysis (PCA) to stochastic processes is the Karhunen-Loève expansion which contributed to solve many problems in the engineering field for more than decade and it seems to be the started point for all the further development of FDA. However in the seventies, Deville (1974), Saporta (1981) and Ramsay (1982) gave the real meaning of the actual approximation to FDA also they provided it by many applications in many fields such as : econometrics, psychology, biology, astronomy and others (see Valderrama (2007)) and we refer to the monographs by : Ramsay and Silverman (1997, 2002 and 2005) also Ferraty and Vieu (2006) for an overview of this topic. According to Ramsay and Silverman (2002 and 2005), the analysis of functional data aims to:

1. formulate the problem at hand in a way amenable to statistical thinking and analysis
2. develop ways of presenting the data that highlight interesting and important features
3. build models for the data observed
4. explain variation in an outcome or dependent variable by using input or independent variable information

In brief, functional data analysis (FDA) aims to model and treat datasets where observations are of functional nature (i.e.FDA has the feature to extend the classical statistical models designed for vectors to the situation when the data are functions or curves (as indicated in Mas (2012))).

2.3 Functional variable and Functional datasets

In many different fields of applied sciences, the collected data are curves which the computing tools, both in terms of memory and computational capacities allow treating a large sets of data. In particular, we can observe a very large set of variables for a single phenomenon (this was pointed out by Ferraty and Vieu (2006)).

To look at the above explanation in another way, a random variable X is called functional variable (f.v) if it takes values in an infinite dimensional space (or functional space). An observation x of X is called a functional data and the observation $x_1, ..., x_n$ of n functional variables $X_1, ..., X_n$ (identically distributed as X) is called a functional dataset.

Obviously, Functional data are a natural extension of multivariate data from finite dimensional to infinite dimensional. "In practice, functional data are obtained by observing a number of subjects over time, space or other continua. The resulting functional data can be curves (including : audiology, biology, environmentology ...), surfaces, or other complex objects"(see Zhang (2014)).

2.4 Functional data : applications on real data

Here, we will try to cover many applied statistics fields by selecting two statistical problems (spectrometric data and phonetic data). For more extensive presentation and a deep explanation, the interested reader can easily refer to : Ramsay and Silverman (2002) and Ferraty and Vieu (2006).

2.4.1 Functional Chemometric data

In many types of chemical problems and for analyzing the chemical composition of any substance, the chemical tools employed to analyze and produced data observation by chemometricians are costly, waste time and sometimes less well known to statisticians in contrary to the spectrometry, that's why; chemometrics appeared to be a field of chemistry that studies the application of statistical methods to chemical data analysis by providing many techniques and several new data-analytical method to the staticians and also the engineers.

When we talk about the problem of quality control in the food industry, it is worth mentioning the first study of Borghard and Tudberg (1992) that used the optimal minimal neural network interpretation of spectra method (OMNIS)which based on principal component analysis as preprocessor to a neural network. The partial least squares and the principal component regression are the most popular methods used to analyze such kind of data.

The following figure plots 215 pieces of finely chopped meat (by random selection), for each curve we represent the observed absorbance as function of the wavelength, these data were obtained from recordings of a Tecator Infractec Food and Feed Analyzer working in the wavelength range of 850-1050 nm by the near-infrared transmission principle. These spectrometric curves are used to predict some chemical properties for instance : the fat content and the percentage of moisture in piece of meat.

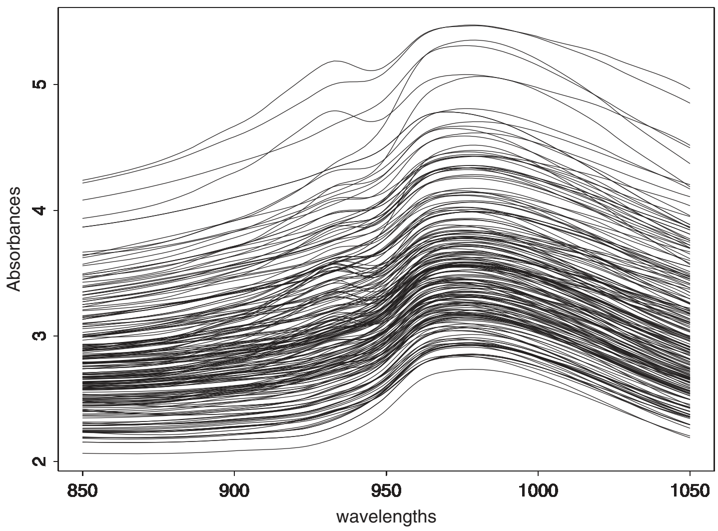

Figure 2.1: The Spectrometric Curves

2.4.2 Speech Recognition data

The observed data in speech recognition have a functional nature, in which they were extracted from the TIMIT database (as pointed out by Zue et al. (1990), “the TIMIT database is the result of a joint effort among researchers at MIT, SRI

International, and Texas Instruments (TI)"). In this study, we concern with five phonemes which are transcribed as follows :

- "sh" as in "she";
- "iy" as the vowel in "she";
- "dcl" as in "dark";
- "aa" as the vowel in "dark";
- "ao" as the first vowel in "water".

The dataset of this study is extracted from 2000 speech frames of 32 ms duration, each speech frame is represented by 400 samples at a 16-kHz sampling rate, from each speech frame, we use log-periodogram as method for casting speech data in a suitable form for speech recognition, our data consist of 2000 log-periodograms of length 150. For more details, we can see : Ferraty and Vieu (2006) (paragraph 2.2).

The following figure shows only 10 log-periodograms curves for each class phoneme.

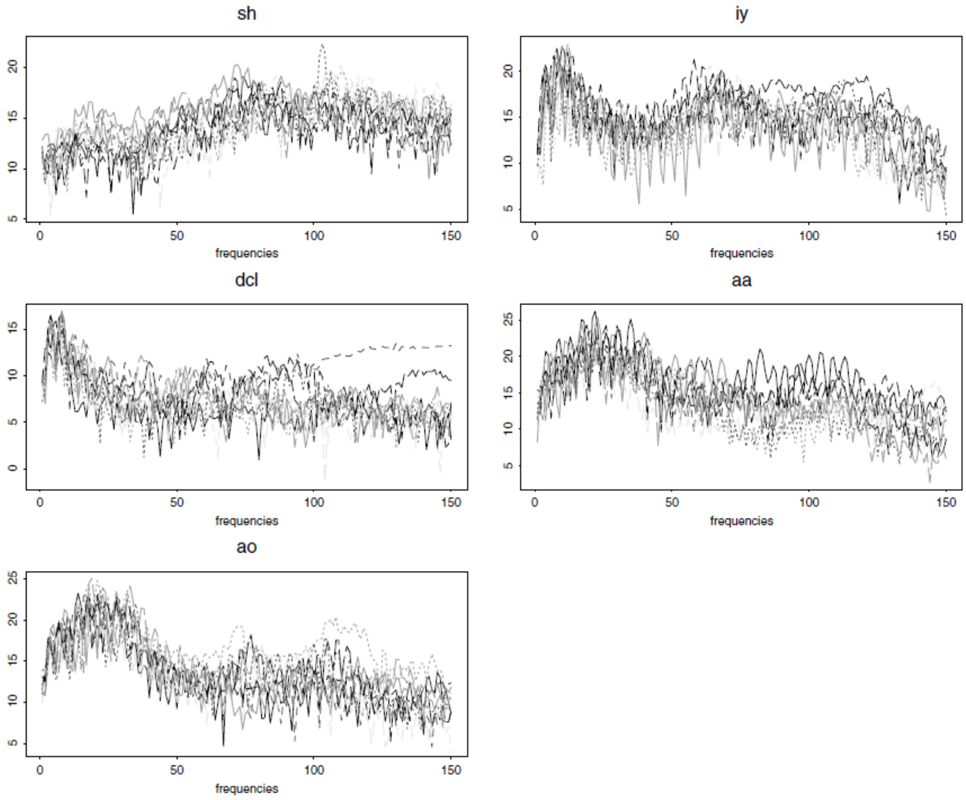

Figure 2.2: A Sample of 10 Log-Periodograms (Curves Data) for each of the Five Phoneme Classes

2.5 Tools and some notations

2.5.1 Semi-metric

During the last decades, distance's notion takes an important place in all statistical methods. Generally, the choice of the metric is not crucial in a finite dimensional space because all the metrics are equivalent rather than in an infinite dimensional space("the equivalence between metrics fails "see Ferraty and Vieu (2006)). In other words, in the functional statistics, the metric spaces can be too restrictive and the choice of the metric becomes crucial. Many studies have postulated the functional variables as random variables with values in $\mathbb{L}^2([0,1])$ (e.g., Crambes et al. (2007)), more extensively in Hilbert space (see Preda (2007), Attaoui *et al.*(2011), Ling and Xu (2012)), Banach (see Cuevas and Fraiman (2004)) or metric (see Dabo-Niang and Rhomari (2003)).

According to madani (2012) (paragraph : 1.2) : in the case of our datasets, it seems that semi-metric spaces are better adapted than metric spaces. Indeed, we can define a semi-metric from a projection of our functional data in low-dimensional spaces either by realizing a functional principal component analysis of our data (Dauxois *et al.* (1982), Hall and Hosseini-Nasab (2006) and Yao and Lee (2006)) or by projecting them on finite basis (wavelets, splines, ...). These reduce the dimension of the data and thus increase the speed of convergence of the methods used while preserving the functional nature of the data. It is well known that the knowledge of the nature of the data function makes it possible to choose the most appropriate basis on which we project our data. For example, we can choose the Fourier basis if we assume that functional variable observed is periodic we refer to the monographs by Ramsay and Silverman (1997, 2005) and Ferraty and Vieu (2006) for an overview on the different approximation methods by projection of functional data and for more discussion about the value of using different types of semi-metric.

2.5.2 Small ball probabilities

The appellation small ball probabilities refers to the behaviour of the smoothing parameter h (also called the bandwidth) when n tends to ∞. In fact, the bandwidth h decreases with the size of the sample of functional variables . More precisely, when we take n very large, h is close to zero and then $B(X, h)$ is considered as a small and $\mathbb{P}(X \in B(X, h))$ is a small ball probability.

In nonparametric statistics for functional variables and precisely in this dissertation, we will show the importance of small ball probabilities in different hypotheses and on rates of convergence. Small ball probabilities are defined by

$$\phi_x(h) = \mathbb{P}(X \in B(X, h)),$$

this function depends on the topological structure existing on the functional semi-metric space and which is induced by the semi-metric itself. When d is a norm, there are many probabilistic results in the literature that study these probabilities of small balls tend to 0 (for instance : Li and Shao (2001) and Gao and Li (2007)) or in the case of seminorm (Aurzada and Simon (2007) and Laksaci (2007)), and for further explanation and discussion about the role of small ball probabilities we can take for instance : Ferraty *et al.*(2006) (paragraph : 5.1) also Ferraty and Vieu (2006) (paragraph : 13.2) as references.

2.5.3 The convergence's notions and Bernstein's inequality

In 1947, the complete convergence conception was introduced by Hsu and Robbins whom proved that the sequence of arithmetic means of i.i.d. random variables (this last have a finite variance) convergences completely to the expected value of the variables and they generalized this result for multivariate random variables. For an updated list of references we refer to : Erdös (1949), Colomb (1984) and Ferraty and Vieu (2006). In this thesis, we will take the almost complete convergence notion as a tool to prove our asymptotic results because this tool is very popular in nonparametric also it implies both the almost sure convergence and the convergence in probability. Therefore, we introduce the following basic definitions

Definition 2.3 *(Ferraty and Vieu (2006)*
Saying $(X_n)_{n\in\mathbb{N}}$ *converges almost completely to some r.r.v.* X*, if and only if*

$$\forall\epsilon > 0, \quad \sum_{n\in\mathbb{N}} \mathbb{P}(|X_n - X| > \epsilon) < \infty,$$

and the almost complete convergence of $(X_n)_{n\in\mathbb{N}}$ *to* X *is denoted by*

$$\lim_{n\to\infty} X_n = X, \ \ a.co.$$

Definition 2.4 *(Ferraty and Vieu (2006))*
we say that the rate of almost complete convergence of $(X_n)_{n\in\mathbb{N}}$ *to* X *is of order* u_n *if and only if*

$$\exists\epsilon_0 > 0, \quad \sum_{n\in\mathbb{N}} \mathbb{P}(|X_n - X| > \epsilon_0 u_n) < \infty,$$

and we write

$$X_n - X = O_{a.co.}(u_n).$$

From many different exponential inequalities and according to our framework study (on functional statistics), we adopt the easiest inequality which is called Bernstein's inequality. Let $Z_1, ..., Z_n$ be independent r.r.v. with zero mean, the next Corollary present the Bernstein's inequality

Corollary 2.1 *(Ferraty and Vieu (2006))*

(i) If $\forall m \geq 2, \exists C_m > 0, \mathbb{E}|Z_1^m| \leq C_m a^{2(m-1)}$*, we have*

$$\forall \epsilon \geq 0, \mathbb{P}\Big(\Big|\sum_{i=1}^{n} Z_i\Big| > \epsilon n\Big) \leq 2\exp\Big\{-\frac{\epsilon^2 n}{2a^2(1+\epsilon)}\Big\}.$$

(ii) Assume that the variables depend on n *(that is, assume that* $Z_i = Z_{i,n}$*). If* $\forall m \geq 2, \exists C_m > 0, \mathbb{E}|Z_1^m| \leq C_m a_n^{2(m-1)}$ *and if* $u_n = n^{-1}a_n^2 \log n$ *verifies* $\lim_{n\to\infty} u_n = 0$*, we have :*

$$\frac{1}{n}\sum_{i=1}^{n} Z_i = O_{a.co.}(\sqrt{u_n}).$$

2.6 Kernel and local linear estimators

In nonparametric and for functional data, this section is devoted to introduce some operators for instance : the regression, the conditional cumulative distribution function and the conditional density function also to provide them by their estimators by using both the kernel method and the local linear method, furthermore we provide these operators by two functional predictors : mode regression and conditional mode.

First of all, we try to clarify our notations by observing n independent pairs of (X_i, Y_i) for $i = 1, ..., n$ identically distributed as (X, Y), this last is valued in $\mathcal{F} \times \mathbb{R}$, where $\mathcal{F}$ is a semi-metric space equipped with a semi-metric d ($d = |\delta|$). Let $S_{\mathcal{F}}$ (resp.$S_{\mathbb{R}}$) be subset of $\mathcal{F}$ (resp.$\mathbb{R}$), and we assume that there exists a regular version of the conditional probability of Y given $X = x$, which is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$.

The regression m of Y on X is defined by :

$$m(x) = \mathbb{E}[Y|X = x],$$

the conditional cumulative distribution function (c.d.f.) of Y given X is defined by :

$$\forall y \in \mathbb{R},\ F^x(y) = \mathbb{P}(Y \leq y|X = x),$$

the conditional density function of Y given X is defined by :

$$\forall y \in \mathbb{R},\ f^x(y) = \frac{\partial}{\partial y}F^x(y).$$

the mode regression function θ on $\mathcal{F}$ is

$$\theta(x) = \sup_{x\in\mathcal{F}} m(x),$$

and the conditional mode $\theta_f(x)$ is defined by :

$$\theta_f(x) = \arg\sup_{y\in S_{\mathbb{R}}} f^x(y)$$

2.6.1 The regression operator

kernel regression estimator

For functional data, Delsol (2008) (paragraph : 3.2.3.1) considered that the linear regression model as the first model was introduced and studied by : Ramsay and Dalzell (1991), Hastie and Mallows (1993). Moreover, the "partial least square" method of regression is generalized to functional case by Preda (1999) then Prada and Saporta (2002, 2004, 2005a and 2005b). In (2000), Ferraty and Vieu introduced the following kernel regression estimator for being a functional extension of Nadaraya-Watson familiar function.

$$\hat{m}(x) = \frac{\sum_{i=1}^{n} Y_i K(h^{-1}||x - X_i||)}{\sum_{i=1}^{n} K(h^{-1}||x - X_i||)},$$

as soon as $||.||$ defines a seminorm
for more asymptotic results and more details about the uses of kernel method we can refer to Ferraty and Vieu (2004, 2006). Burba et al (2009) investigated the k-nearest-neighbors(k-NN) estimator. Dabo-Niang and Rhomani (2003) : convergence in $\mathbb{L}^p$ norm, the asymptotic normality in the α-mixing case(see Masry (2005)) and Delsol (2007a, 2007b) is generalized the results in Dabo-Niang and Rhomani (2003) to dependent samples.

Local linear regression estimator

local linear regression has drawn much attention and date back to more than 20 years (for more details see : Wand and Jones (1995) for univariate or multivariate explanatory variables). local linear regression models have been carried out when the explanatory variable is a functional predictor by Boj *et al.* (2008) and Aneiros-Pérez *et al.* (2011) and when the explanatory variable takes values in a Hilbert space, the $\mathbb{L}^2$-convergence rate of the local linear estimate of the regression function is stated by Baíllo and Grané (2009). However, in a semi-metric space, Barrientos-Marin*et al.* (2010) propose an alternative simplified version of the local linear estimator of the regression operator in the i.i.d. setting and they establish the rate of the almost complete convergence of the proposed estimate. In the remainder of this thesis, we adopt the locally modelled regression estimator of Barrientos-Marin et al. (2010). Therefore, to construct a functional local linear estimator of the regression operator, we consider the following minimization problem :

$$\min_{(a,b)\in\mathbb{R}^2} \sum_{i=1}^{n} (Y_i - a - b\beta(X_i, x))^2 K\left(\frac{|\delta(X_i, x)|}{h_K}\right)$$

where : $\beta(.,.)$ and $\delta(.,.)$ are two functions defined from $\mathcal{F} \times \mathcal{F}$ to $\mathbb{R}$, such that: $\forall \xi \in \mathcal{F}, \beta(\xi,\xi) = 0,$ and $d(.,.) = |\delta(.,.)|$. K is Kernel and $h_K = h_{K,n}$ is chosen as a sequence of positive real numbers which converges to 0 when $n \to \infty$. Here we denote : $\hat{a}$ by $\hat{m}$ and $\hat{b}$ by $\hat{m}^{(1)}$ (as the estimator of the first order derivative of m). By simple algebra, we get the following explicit solution :

$$\hat{m}(x) = \frac{\sum_{i,j=1}^{n} W_{i,j}(x) Y_j}{\sum_{i,j=1}^{n} W_{i,j}(x)}$$

whith

$$W_{i,j}(x) = \beta(X_i, x)(\beta(X_i, x) - \beta(X_j, x))K(h_K^{-1}\delta(X_i, x))K(h_K^{-1}\delta(X_j, x))$$

More recently, findings have emerged that offers many asymptotic properties like, the asymptotic normality of the functional local linear regression estimate (see Zhou and Lin (2015)) and the strong convergence (with rates) uniformly in bandwidth (UIB) consistency of local linear regression estimate (see Attouch *et al.* (2018)), just to name a few.

2.6.2 The conditional cumulative distribution function

Kernel estimator of conditional cumulative distribution function

In the last two decades and in infinite dimensional spaces, the behaviour of kernel type estimation of the conditional cumulative distribution function had received a lot of attention . In (2006), Ferraty *et al.* have introduced the kernel type estimation of some characteristics of the conditional cumulative distribution function also the successive derivatives of the conditional density and the asymptotic properties are established for each of these estimates. Recently, the uniform in bandwidth consistency of conditional distribution kernel estimator has been proved by Kara-Zaitri et al.(2017).

The kernel conditional c.d.f. estimator is defined as follows :

$$\hat{F}^x(y) = \frac{\sum_{i=1}^{n} K(h^{-1}d(x, X_i))H(h_H^{-1}(y - Y_i))}{\sum_{i=1}^{n} K(h^{-1}d(x, X_i))},$$

where : H is a c.d.f and h_H is a strictly positive real number (depending on n)

Local linear estimator of conditional cumulative distribution function

The trick of obtaining the estimator of conditional cumulative distribution function by local linear method is to take the function $F^x(.)$ as a regression model with the response variable $H(h_H^{-1}(. - Y_i))$, where H is some cumulative distribution function and ($h_H = h_{H,n}$)is a sequence of positive real numbers (see Fan and Gijbels (1996)and Demongeot et al.(2014)). we can summarize this idea by the following motivation :

$$\mathbb{E}[H(h_H^{-1}(y - Y_i))|X_i = x] \to F^x(y) \quad \text{as} \quad h_H \to 0$$

the construction of functional local linear estimator of the conditional distribution function was based on the minimization of the following equation

$$\min_{(a,b)\in\mathbb{R}^2} \sum_{i=1}^{n} (H(h_H^{-1}(y - Y_i)) - a - b\beta(X_i, x))^2 K\left(\frac{\delta(X_i, x)}{h}\right)$$

then, our estimator is explicitly defined by :

$$\hat{F}^x(y) = \frac{\sum_{i=1}^{n} W_{i,j}(x)H(h_H^{-1}(y - Y_i))}{\sum_{i=1}^{n} W_{i,j}(x)},$$

where

$$W_{i,j}(x) = \beta(X_i, x)(\beta(X_i, x) - \beta(X_j, x))K(h_K^{-1}\delta(X_i, x))K(h_K^{-1}\delta(X_j, x)).$$

In this context, Demongeot et al.(2014) stated the almost-complete and L^2-consistency of $\hat{F}^x(y)$ (with rates) and they invested their asymptotic results to discuss some statistical problems such as the choice of the smoothing parameters and the determination of confidence intervals. Until lately, Bouanani et al.(2019) considered the asymptotic normality of the local linear estimator of the conditional cumulative distribution in the i.i.d. setting.

2.6.3 The conditional density function

Kernel estimator of conditional density function

In statistical analysis for functional data and in the independent and identically distributed (i.i.d.) case, the first detailed study of conditional density and its derivatives was introduced by Ferraty et al. (2006). They established the asymptotic properties (the almost complete consistency) of the kernel estimator of the conditional density function and the asymptotic normality was stated by Ezzahrioui and Ould Saïd (2008).These results have been extended to dependent data by Ezzahrioui and Ould Saïd (2010). The kernel type estimator of conditional density is introduced by

$$\hat{f}^x(y) = \frac{h_H^{-1}\sum_{i=1}^{n} K(h^{-1}d(x, X_i))H(h_H^{-1}(y - Y_i))}{\sum_{i=1}^{n} K(h^{-1}d(x, X_i))}$$

this kernel model has been employed to establish the rate of the uniform almost complete convergence (see Ferraty et al.(2010)).Furthermore, the previous estimator has been extended when the observations are linked with a single-index structure such as the pointwise and the uniform almost complete convergence (with the rate) (see Attaoui et al.(2011)) and in the α-mixing functional data (see Ling and Xu (2012)).

Local estimator of conditional density function

In this part, we adopt the fast functional local modelling by minimizing the following quantity to estimate the conditional density of a scalar response variable given a random variable taking values in a semi-metric space

$$\min_{(a,b)\in\mathbb{R}^2} \sum_{i=1}^{n} (h_H^{-1} H(h_H^{-1}(y - Y_i)) - a - b\beta(X_i, x))^2 K\left(\frac{\delta(X_i, x)}{h}\right)$$

and we get the following expression :

$$\hat{f}^x(y) = \frac{\sum_{i=1}^{n} W_{i,j}(x) H(h_H^{-1}(y - Y_i))}{h_H \sum_{i=1}^{n} W_{i,j}(x)},$$

where

$$W_{i,j}(x) = \beta(X_i, x)(\beta(X_i, x) - \beta(X_j, x))K(h_K^{-1}\delta(X_i, x))K(h_K^{-1}\delta(X_j, x))$$

this model was introduced by Demongeot et al.(2013) also they established both the pointwise and the uniform almost-complete consistencies with convergence rates of the conditional density estimator.Moreover, the quadratic error in the local linear estimation of the conditional density was introduced by Rachdi et al. (2014)(with some comments and discussions like bandwidths selection and confidence intervals.), recently, the joint asymptotic normality of the estimators of the conditional density and its derivative is established in the α-mixing setup by Xiong et al. (2018) and in the i.i.d. setting by Bouanani et al.(2019).

References

ANEIROS-PÉREZ, G.,CAO, R. and VILAR-FERNÁNDEZ, J.M. (2011). Functional methods for time series prediction : A nonparametric approach. *J of Forecasting, **30**, 377-392.*

ATTAOUI, S., LAKSACI, A. *and* OULD-SAÏD, E. *(2011). A note on the conditional density estimate in the single functional index model.* Statist. Probab. Lett, **81**(1), 45-53.

ATTOUCH, M., LAKSACI, A. and RAFAA, F.(2018). On the local linear estimate for functional regression : Uniform in bandwidth consistency. *Communications in Statistics-Theory and Methods, **48**, 1-18.*

AURZADA, F. *and* SIMON, T. *(2007). small ball probabilities for stable convolutions.* ESAIM Probab. Stat, **11**, 327-343 (electronic).

BAÍLLO, A. and GRANÉ, A. (2009). Functional Local Linear Regression with Functional Predictor and Scalar Response. *Journal of Multivariate Analysis, **100**, 102-111.*

BARRIENTOS-MARIN, J., FERRATY, F. *and* VIEU, P. *(2010). Locally Modelled Regression and Functional Data.* J. of Nonparametric Statistics ***22**, 617-632.*

BOJ, E. DELICADO, P. *and* FORTIANA, J. *(2010). Distance-based local linear regression for functional predictors.* Computational Statistics and Data Analysis, **54**, 429-437.

BORGGAARD, C., THODBERG, H.H.(1992). Optimal Minimal Neural Interpretation of Spectra. Analytical Chemistry, **64**, 545-551.

BOUANANI, O., LAKSACI, A., RACHDI, M. and RAHMANI, S. (2019). Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics. *Behaviormetrika, **46**, 199-233.*

BURBA, F.;FERRATY, F. *and* VIEU, P. *(2009). k-Nearest Neighbour method in functional nonparametric regression, Journal of Nonparametric Statistics 21(4), 453-469 .*

COLLOMB, G.*(1984). Propriétés de convergence presque compléte du prédicteur à noyau (in french).*Z. Wahrscheinlichkeitstheorie verw. Gebiete, **66**, 441-460.

CRAMBES, C., KNEIP, A. and SARDA, P. (2007). Smoothing splines estimators for functional linear regression, *Ann. Statist, **37**(1), 35-72.*

CUEVAS, A. *and* FRAIMAN, R. *(2004). On the bootstrap methodology for functional data,* COMPSTAT 2004-Proceedings in Computational Statistics, 127-135.

DABO-NIANG, S. and RHOMARI, N. (2003). Kernel regression estimation when the regressor takes values in metric space. *C. R. Acad. Sci. Paris*, **336**, 75-80.

DAUXOIS, J., POUSSE, , A. and ROMAIN, Y. (1982). Asymptotic theory for the principal component analysis of a vector random function : some applications to statistical inference, *J. Multivariate Anal, **12**(1), 136-154.*

DELSOL, L. *(2007a). CLT and $\mathbb{L}^q$ errors in nonparametric functional regression,* C. R. Math. Acad. Sci, **345** (7), 411-414.

DELSOL, L. (2007b). Régression non-paramétrique fonctionnelle : Expressions asymptotiques des moments, *Annales de l'I.S.U.P, **LI**(3), 43-67.*

DELSOL, L. *(2008). Régression sur variable fonctionnelle : Estimation, Tests de structure et Applications. PhD thesis from the Toulouse III - Paul Sabatier University (France).* DEMONGEOT, J., LAKSACI, A., MADANI, F. *and* RACHDI, M. *(2013). Functional data : local linear estimation of the conditional density and its application.*Statistics*, **47**, 26-44.*

DEMONGEOT, J., LAKSACI, A. RACHDI, M. *and* RAHMANI, S. *(2014). On the local Modalization of the conditional distribution for functional data.* Sankhya A*, **76** (2), 328-355.*

DEVILLE, JC. *(1974). Méthodes statistiques et numériques de l'analyse harmonique.* Ann INSEE, ***15**, 3-101.*

ERDÖS, P. *(1949). On a theorem of Hsu and Robbins.* Ann. Math. Statist, **20** (2), 286-291.

EZZAHRIOUI, M. and OULD-SAÏD,E. (2008). Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *J Nonparametric Stat, **20**, 3-18.*

EZZAHRIOUI, M. *and* OULD SAÏD, E. *(2010).Some asymptotic results of a nonparametric conditional mode estimator for functional time series data.* Neerlandica, **64**, 171-201.

FAN, J. GIJBELS, I. (1996). *Local Polynomial Modelling and its Applications.* Chapman & Hall, London.

FERRATY, F. and VIEU, P. (2000). Dimension fracale et estimation de la régression dans des espaces vectoriels semi-normés, *Compte Rendus de l'Académie des Sciences Paris, **330**, 403-406.*

FERRATY, F. *and* VIEU, P. *(2004). Nonparametric models for functional data, with application in regression, time series prediction and curve discrimination.* J. Nonparametric. Stat. **16** (1-2) 111-125.

FERRATY, F.,LAKSACI, A. andVIEU, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models. *Statistical Inference for Stochastic Processes*, **9**, 47-76.

FERRATY, F. and VIEU, P. (2006). *Nonparametric functional data analysis : Theory and Practice. Springer Series in Statistics, New York : Springer-Verlag.*

FERRATY, F., LAKSACI, A., TADJ, A., *and* VIEU, P.*(2010). Rate of uniform consistency for nonparametric estimates with functional variables.* Journal of Statistical Planning and Inference*, **140**, 335-352.*

GAO, F. *and* LI, W.V. *(2007). Small ball probabilities for the Slepian Gaussian fields.* Transactions of the American Mathematical Society, **359** (3), 1339-1350. (electronic)

HALL, P. and HOSSEINI-NASAB, M. (2006). On properties of functional principal components. *Journal of the Royal Statistical Society Series B, **68**, 109-126.*

HASTIE, T. *and* MALLOWS, C. *(1993) Discussion of "A statistical view of some chemometrics regression tools."by Frank, I.E. and Friedman,* J.H. Technometrics, **35**, 140-143.

HSU, P.and ROBBINS, H. (1947). Complete convergence and the law of large numbers. *Proc. Nat. Acad. Sci. USA, **33**, 25-31.*

KARA, L-Z., LAKSACI, A., RACHDI, M. *and* VIEU, P. *(2017). Uniform in bandwidth consistency for various kernel estimators involving functional data.* Journal of Nonparametric Statistics, **29**, 85-107.

LAKSACI, A. (2007). Erreur quadratique de l'estimateur à noyau de la densité conditionnelle à variable explicative fonctionnelle. *C. R. Acad. Sci. Paris,***345**, 171-175.

LI, W.V. and SHAO, Q.M. (2001). Gaussian processes : inequalities, small ball probabilities and applications, In :C.R. Rao and D. Shanbhag (eds.) Stochastic processes, Theory and Methods. Handbook of Statitics, 19, North-Holland, Amsterdam.

LING, N. and XU, Q. (2012). Asymptotic normality of conditional density estimation in the single index model for functional time series data. *Statistics and Probability Letters*, **82**, 2235-2243.

MADANI, F. (2012). Aspects théoriques et pratiques dans l'estimation non paramétrique de la densité conditionnelle pour des données fonctionnelles. PhD thesis from the Grenoble University (France).

MANTEIGAA, W. G. and VIEU, P. (2007). Editorial statistcs for functional data. *Computational Statistics & Data Analysis, **51**, 4788-4792.*

MAS, A. *(2012). Lower bound in regression for functional data by representation of small ball probabilities.* Electronic Journal of Statistics, ***6**, 1745-1778.*

MASRY, E.*(2005). Nonparametric regression estimation for dependent functional data : asymptotic normality. Stochastic Process and their Applications, **115** (1), 155-177.*

PREDA, C. *(1999). Analyse factorielle d'un processus : problème d'approximation et de régression (in french), PhD Lille I, 1999.*

PREDA, C. *and* SAPORTA, G. *(2002). Régression PLS sur un processus stochastique. Revue de Statistique Appliquée, **50** (2), 27-45.*

PREDA, C. *and* SAPORTA, G. *(2004). PLS approach for clusterwise linear regression on functional data. Classification, clustering, and data mining applications. 167-176, Stud. Classification Data Anal. Knowledge Organ., Springer, Berlin.*

PREDA, C. *and* SAPORTA, G. *(2005a). PLS regression on a stochastic process.Comput.Statist. Data Anal, **48** (1), 149-158.*

PREDA, C. *and* SAPORTA, G.*(2005b). Clusterwise PLS regression on a stochastic process. Comput.Statist. Data Anal,**49** (1), 99-108.*

PREDA, C. *(2007). Regression models for functional data by reproducing kernel Hilbert spaces methods.* J. Statist. Plann. Inference, **137**(3), 829-840.

RACHDI, M., LAKSACI, A.,DEMONGEOT, J., ABDALI, A. and MADANI, F. (2014). Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data. *Computational Statistics and Data Analysis*, **73**, 53-68.

RAMSAY, J. O. (1982). When data are functions. *Psychometrika, **47**(4), 379-396.*

RAMSAY, J. *and* DALZELL, C. *(1991). Some tools for functional data analysis.* J. R. Statist. Soc. B, **53**, 539-572.

RAMSAY, J. O. and SILVERMAN, B.W.(1997). Functional data analysis. Springer series in statistics. Springer, New York.

RAMSAY, J. O. and SILVERMAN, B.W. (2002). Applied functional data analysis: Methods and Case Studies. Springer Series in Statistics. Springer-Verlag, New York.

RAMSAY, J. O. and SILVERMAN, B.W. (2005). Functional Data Analysis. 2nd ed.Springer, New-York.

SAPORTA, G. (1981). Méthodes exploratoires d'analyse de données temporelles. *Cahiers du Bureau universitaire de Recherche Opérationnelle Série Recherche,* ***37****(38), 7-194.*

VALDERRAMA MJ. *(2007).* An overview to modelling functional data. *Computational Statistics,* ***22****, 331-334.*

WAND, M. *and* JONES, C.*(1995).* Kernel Smoothing, Monographs on Statistics and Applied Probability (Vol. 60), London : Chapman & Hall.

XIONG, X., ZHOU, P. and AILIAN, C. (2018). Asymptotic normality of the local linear estimation of the conditional density for functional time series data. *Communications in Statistics-Theory and Methods,* ***47****(14), 3418-3440.*

YAO, F. *and* LEE, T.C.M. *(2006). Penalised spline models for functional principal component analysis,* Journal of the Royal Statistical Society Series B, **68** (1), 3-25.

ZHANG, J.T. (2014). *Analysis of Variance for Functional Data.* New York: Chapman and Hall/CRC.

ZHOU, Z. and LIN, Z. (2015). Asymptotic normality of locally modelled regression estimator for functional data, Journal of Nonparametric Statistics, **28**, 1-16.

Zue, V., Seneff, S., Glass, J. (1990). Speech database development at MIT: TIMIT and BEYOND. Speech Communication, **9**, 351-356

Chapter 3

Some asymptotic results linked with nonparametric estimation for functional covariates

To date, the asymptotic properties take an important attention to clarify the asymptotic behaviour of many model estimates. The current chapter is mainly focuses on nonparametric functional data framework in which we attempt to summarize some results of many papers for each of : regression, conditional cumulative distribution function and conditional density, also we provide them by their stochastic mode of convergence, this latter comes under technical assumptions.

3.1 Regression

To update our information and give a right image of regression, this section is devoted to give some important results for both methods : kernel and local linear which help us to describe the asymptotic behaviour of our operator.

3.1.1 Kernel regression

Through this paragraph, we adopt the following kernel regression estimator of the regression operator m

$$\hat{m}(x) = \frac{\sum_{i=1}^{n} Y_i K(h^{-1}d(x, X_i))}{\sum_{i=1}^{n} K(h^{-1}d(x, X_i))},$$

for making the previous model more flexible and easy to treat, Ferraty and Vieu(2006) putted forward the next technical conditions
Assumptions and asymptotic behaviour :

(H1) for all $\epsilon > 0,\ \mathbb{P}(X \in B(x, \epsilon)) = \phi_x(\epsilon) > 0.$
To introduce the pointewise almost complete convergence, we need the following continuity-type hypothesis :

(H2$_C$) $m \in \left\{ f : \mathcal{F} \to \mathbb{R}, \lim_{|\delta(x,x')| \to 0} f(x') = f(x) \right\},$
furthermore, the next Lipschitz-type constraint allows us to find the rate of convergence

(H2$_L$) $m \in \left\{ f : \mathcal{F} \to \mathbb{R}, \exists C \in \mathbb{R}_+, x' \in \mathcal{F}, |f(x) - f(x')| < C|\delta(x, x')|^\beta \right\}$, where $\beta \geq 0$.

(H3) the bandwidth h is a positive sequence such that
$\lim_{n\to\infty} h = 0$ and $\lim_{n\to\infty} \log n / n\phi_x(h) = 0$

(H4) K is a kernel of type I
or K is a kernel of type II and $\phi_x(.)$ satisfies

$$\exists C_3 > 0,\ \exists \epsilon_0,\ \forall \epsilon < \epsilon_0,\ \int_0^\epsilon \phi_x(u)du > C_3 \epsilon \phi_x(\epsilon).$$

(H5) $\forall m \geq 2, \mathbb{E}[|Y^m||X = x] < \sigma_m(x) < \infty$ with $\sigma_m(.)$ continuous at x.

Theorem 3.1 *Ferraty and Vieu(2006)*
under (H1), ($H2_C$) and (H3)-(H5), we get

$$\lim_{n\to\infty} \hat{m}(x) = m(x), \, a.co.$$

Theorem 3.2 *Ferraty and Vieu(2006)*
Under assumptions (H1), ($H2_L$) and (H3)-(H5), we obtain

$$\hat{m}(x) - m(x) = O(h^{\beta}) + O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right).$$

By using an easy changes in our notations like $B^x_{h_n}$ is the closed ball and μ is the law of the functional variable X, for $p \geq 1$ and under general assumptions, the next paragraph is tackled to study the p-mean convergence (of $\hat{m}$) that is realized by Dabo-Niang and Rhomari (2003).

Assumptions and asymptotic behaviour :

(M1) There exists r, a and $b > 0$ such that

$$a\mathbb{1}_{\{|u|\leq r\}} \leq K(u) \leq b\mathbb{1}_{\{|u|\leq r\}}.$$

(M2) $h_n \to 0$ and $\lim_{n\to\infty} n\mu(B^x_{rh_n}) = \infty$,

(M3) $\lim_{h\to 0} \frac{1}{\mu(B^x_h)} \int_{B^x_h} |m(w) - m(x)|^p d\mu(w) = 0.$

(M4) $\mathbb{E}\left[|Y - m(X)|^p | X \in B^x_{rh_n}\right] = o([n\mu(B^x_{rh_n})]^{p/2}).$

(M5) We suppose that m is "p-mean Lipschitzian ", of parameters $0 < \tau = \tau_x \leq 1$ and $c_x > 0$, in a neighborhood of x :

$$\frac{1}{\mu(B^x_h)} \int_{B^x_h} |m(w) - m(x)|^p d\mu(w) \leq c_x h^{p\tau}, \text{ when } h \to 0.$$

Theorem 3.3 *the p-mean convergence (Dabo-Niang and Rhomari (2003))*
(i) under (M1)-(M4) and if $\mathbb{E}(|Y|^p) < \infty$*, we have*

$$\mathbb{E}(|\hat{m}(x) - m(x)|^p) \xrightarrow[n\to\infty]{} 0$$

(ii) In addition, if $\mathbb{E}\left[|Y - m(X)|^p | X \in B^x_{rh_n}\right] = O(1)$, $\mathbb{E}(|Y|^p) < \infty$ *with* $p \geq 2$ *and (M1), (M2) and (M5) are fulfilled, then we have*

$$\mathbb{E}\left(|\hat{m}(x) - m(x)|^p\right) = O\left(h_n^{p\tau} + \left(\frac{1}{n\mu(B^x_{rh_n})}\right)^{p/2}\right).$$

3.1.2 Local linear regression

In Hilbertian space, Baíllo and Grané (2009) considered the regression problem with an explanatory functional variable X taking values in $L^2[0,1]$ and a scalar response Y. Let $(X_i, Y_i)_{i=1,\dots,n}$ be n independent pairs, identically distributed as (X, Y). Baíllo and Grané (2009) presented the following minimization to get $\hat{m}$:

$$\sum_{i=1}^{n} (Y_i - (a + < b, X_i - x >))^2 K_h(||X_i - x||). \tag{3.1}$$

Where : $K_h(.) := h^{-1}K(./h), h = h_n, b = b(x) \in L^2[0,1]$and $<,>$ is the $L^2[0,1]$ inner product.
Once the value $\hat{a}$ of a minimizing (3.1) is the local linear estimator $(\hat{m}_{LL}(x))$ of $m(x)$.
In order to reduce the dimension of parameter b, they used an orthonormal basis $\{\phi_j\}_{j\geq 1}$ of $L^2[0,1]$,

$$b = \sum_{j=1}^{J} b_j \phi_j \quad \text{and} \quad X_i - x = \sum_{j=1}^{J} c_{ij}\phi_j$$

with

$$b_j =< b, \phi_j > \quad \text{and} \quad c_{ij} =< X_i - x, \phi_j > .$$

the local linear estimator of $m(x)$ is :

$$\hat{m}_{LL}(x) = \hat{a} =^t \boldsymbol{e}_1({}^t\boldsymbol{CWC})^{-1}\ {}^t\boldsymbol{CWY} \tag{3.2}$$

where : $\boldsymbol{e}_1$ is the $(J+1) \times 1$ vector having 1 in the first entry and 0's in the rest, $\boldsymbol{Y} =^t (Y_1, ..., Y_n)$, $\boldsymbol{W} = \text{diag}(K_h(||X_1 - x||), ..., K_h(||X_n - x||))$ and

$$\boldsymbol{C} = \begin{pmatrix} 1 & c_{11} & \dots & c_{1J} \\ 1 & c_{21} & \dots & c_{2J} \\ \vdots & & \vdots & \\ 1 & c_{n1} & \dots & c_{nJ} \end{pmatrix}.$$

In order to study the asymptotic behaviour of $\hat{m}_{LL}(x)$ and to establish the next theorem, Baíllo and Grané (2009) introduced the following conditions
Assumptions and asymptotic behaviour :

(A1) The kernel $K : \mathbb{R} \to \mathbb{R}^+$ satisfying $\int K = 1$ is a kernel of type I

(A2) $\forall \epsilon > 0$, the small ball probability $\phi_x(.)$ satisfies $\phi_x(\epsilon) := \mathbb{P}(||X - x|| < \epsilon) > 0$.

(A3) With probability one, any trajectory $X(., \omega)$ of X has derivative of ν-th order which is uniformly bounded on $[0, 1]$ by a constant independent of ω.

(A4) The ν-th order derivative of the element x exists and is uniformly bounded on $[0, 1]$.

(A5) The regression function m is differentiable in a neighbourhood of x and twice differentiable at x with continuous second derivative,and also the bandwidth h satisfies :
$h \to 0$ and $\lim_{n\to\infty} n\phi_x(h) = \infty$

Theorem 3.4 *Baíllo and Grané (2009)*
Let the assumptions (A1)-(A5) hold, then

$$\mathbb{E}[(\hat{m}_{LL}(x) - m(x))^2 | X] = \left(O(J^{-\nu}) + O_p(h^2)\right)^2 + O_p\left((n\phi_x(h))^{-1}\right).$$

However, in a semi-metric space, Barrientos et al.(2010) putted forward another alternative version of the local linear estimator of the regression operator (as pointed out in previous chapter paragraph : 1.6.1).

Assumptions and asymptotic results

In the i.i.d. setting, Barrientos et al.(2010) stated the pointewise almost complete convergence and the rate of convergence of the proposed estimate, for this reason, they gave these a crucial hypotheses which are :

(H1) $\phi_x(u_1, u_2) := \mathbb{P}(u_1 \leq \delta(X, x) \leq u_2)$, and $\forall u > 0, \phi_x(u) := \phi_x(0, u) > 0$.
Where : $\phi_x(u)$ is the probability of a ball of $\mathcal{F}$ centered at x and of radius u, when u tends to 0, it is clear that $\phi_x(u)$ is called small ball probability function.
To introduce the rate of pointewise almost complete convergence, we need the following Lipschitz-type constraint

(H2) $m \in \{f : \mathcal{F} \to \mathbb{R}, \exists C \in \mathbb{R}_+, x' \in \mathcal{F}, |f(x) - f(x')| < C|\delta(x, x')|^v\}$, where $v \geq 0$.

(H3) there exists $0 < M_1 < M_2$ and $\forall x' \in \mathcal{F}$,

$$M_1|\delta(x, x')| \leq |\beta(x, x')| \leq M_2|\delta(x, x')|.$$

(H4) The kernel function K : is a positive function and differentiable on its support$[0, 1]$.

(H5) $\lim_{n\to\infty} h = 0$ and $\lim_{n\to\infty} \log n/n\phi_x(h) = 0$.

(H6) The behaviour of the bandwidth h is :

$$\exists n_0, \forall n > n_0, \frac{1}{\phi_x(h)} \int_0^1 \phi_x(zh, h)\frac{d}{dz}(z^2K(z))dz > C > 0$$

(H7) The local expectation of β satisfies : $h \int_{B(x,h)} \beta(u, x)dP(u) = o\Big(\int_{B(x,h)} \beta^2(u, x)dP(u)\Big)$.

(H8) For all $k \geq 2, \sigma_k : x \to \mathbb{E}[Y^k|X = x]$ is a continuous operator.

Theorem 3.5 *Barrientos et al.(2010)*
Under (H1)-(H8), we get

$$\hat{m}(x) - m(x) = O(h^v) + O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right)$$

In the same framework of locally modelled regression estimator for functional data, Zhou and Lin (2015) established the mean-squared convergence and asymptotic normality for the estimator. Hence, they adopted the following regression estimator :

$$\hat{m}(x) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}Y_j}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} = \frac{\sum_{j=1}^{n} W_j K_j Y_j}{\sum_{j=1}^{n} W_j K_j},$$

with

$w_{ij} = \beta_i(\beta_i - \beta_j)K_iK_j, W_j = \sum_{i=1}^{n}(w_{ij}/K_j) = \sum_{i=1}^{n}\beta_i^2 K_i - (\sum_{i=1}^{n}\beta_i K_i)\beta_j$

Assumptions and asymptotic results

for any fixed $x \in \mathcal{F}$, Zhou and Lin (2015) imposed these conditions :

(H.1) $\forall r > 0, \phi_x(r) = \mathbb{P}(|\delta(X, x)| \leq r)$. Furthermore, there exists a function $\Phi_x(u)$ such that

$$\lim_{h \to 0} \frac{\phi_x(uh)}{\phi_x(h)} = \Phi_x(u), \forall u \in [0, 1].$$

(H.2) m and σ^2 are both in the set

$$\left\{ f : \mathcal{F} \to \mathbb{R}, \lim_{|\delta(x',x)| \to 0} f(x') = f(x) \right\}.$$

(H.3) Denoting that $B(x, r) = \{x' \in \mathcal{F} : |\delta(x, x')| \leq r\}$ and $P_x(z)$ is the probability distribution of x. The bi-functional operator β is such that :

(H.3.1) $\exists 0 < C_1 < C_2, \forall x' \in \mathcal{F}, C_1|\delta(x, x')| \leq |\beta(x, x')| \leq C_2|\delta(x, x')|$;

(H.3.2) $\sup_{x \in B(x,r)} |\beta(u, x) - \delta(u, x)| = o(r)$;

(H.3.3) $h \int_{B(x,h)} \beta(u, x) dP_x(u) = o(\int_{B(x,h)} \beta^2(u, x) dP_x(u))$.

(H.4) The kernel K is supported on $[0, 1]$ and has a continuous derivative $K'(s) < 0$ for $s \in [0, 1)$ and $K(1) > 0$.

(H.5) $\lim_{n \to \infty} h = 0$ and $\lim_{n \to \infty} n\phi_x(h) = \infty$.

Before giving the asymptotic results, authors gave us some notations; so, they denoted that:

$$M_j = K^j(1) - \int_0^1 (K^j(u))'\Phi_x(u)du, \text{ where } j = 1, 2;$$

$$N(a, b) = K^a(1) - \int_0^1 (u^b K^a(u))'\Phi_x(u)du, \ \forall a > 0 \text{ and } b = 2, 4;$$

on the other hand , we can take $\hat{m}$ as :

$$\hat{m}(x) = \hat{m}_1(x)/\hat{m}_0(x) \text{ with } \hat{m}_l(x) = \frac{1}{n\mathbb{E}(W_1K_1)}\sum_{j=1}^{n} W_jK_jY_j^l, \text{for } \ l = 0, 1;$$

and

$$\begin{aligned} B_n(x) = \mathbb{E}[\hat{m}_1(x)] - m(x) &= \frac{\mathbb{E}(W_1K_1Y_1)}{\mathbb{E}(W_1K_1)} - m(x) \\ &= \frac{\mathbb{E}(\beta_1^2K_1).\mathbb{E}[K_1(m(x_1) - m(x))] - \mathbb{E}(\beta_1K_1).\mathbb{E}[\beta_1K_1(m(x_1) - m(x))]}{\mathbb{E}(\beta_1^2K_1).\mathbb{E}(K_1) - \mathbb{E}^2(\beta_1K_1)}. \end{aligned}$$

Theorem 3.6 *Mean-squared convergence (Zhou and Lin (2015))*
under (H.1)-(H.5), we have that

$$\mathbb{E}[\hat{m}(x)] - m(x) = B_n(x) + O\left(\frac{1}{n\phi_x(h)}\right)$$

and

$$Var[\hat{m}(x)] = \frac{1}{n\phi_x(h)}\frac{M_2}{M_1^2}\sigma^2(x) + o\left(\frac{1}{n\phi_x(h)}\right)$$

Theorem 3.7 *Asymptotic normality (Zhou and Lin (2015))*
under (H.1)-(H.5), we have that

$$\sqrt{n\phi_x(h)}(\hat{m}(x) - m(x) - B_n(x)) \xrightarrow{d} \mathcal{N}\left(0, \frac{M_2}{M_1^2}\sigma^2(x)\right).$$

3.2 Conditional distribution function

To open more our appetite about the uses of kernel method and local method, this section provides an overview about the behaviour of conditional cumulative distribution function which has great statistical merits.

3.2.1 Kernel conditional distribution function estimator

We mainly spotlight on the study of Ferraty et al.(2006) to introduce the rate of pointwise almost complete convergence of the functional kernel estimator of the conditional cumulative distribution function $F^x(y)$, which was defined in previous chapter paragraph (1.6.2).
For $x \in \mathcal{F}$ and $\forall y \in \mathbb{R}$, the conditional cumulative distribution function of Y given $X = x$ is defined by :

$$F^x(y) = \mathbb{P}(Y \leq y|X = x),$$

moreover, we assume that F^x is absolutely continuous with respect to the Lebesgues measure on $\mathbb{R}$.

Hereinafter, we denote by : N_x is a fixed neighborhood of x and $S_{\mathbb{R}}$ is a fixed compact subset of $\mathbb{R}$ and we introduce the following conditions to give the rate of pointwise almost complete convergence of our estimator (see Ferraty et al.(2006)).

Assumptions and asymptotic results

(H1) for all $\epsilon > 0, \mathbb{P}(X \in B(x, \epsilon)) = \phi_x(\epsilon) > 0$.

(H2) $\forall (x_1, x_2) \in N_x^2, \forall (y_1, y_2) \in S_{\mathbb{R}}^2$,

$$|F^{x_1}(y_1) - F^{x_2}(y_2)| \leq C_x(d(x_1, x_2)^{b_1} + |y_1 - y_2|^{b_2})$$

(H3) $\begin{cases} \forall (y_1, y_2) \in \mathbb{R}^2, |H(y_1) - H(y_2)| \leq C|y_1 - y_2|, \\ \int |t|^{b_2} H^{(1)}(t)dt < \infty. \end{cases}$

(H4) The kernel K is a function with support $(0, 1)$ such that

$$0 < C_1 < K(t) < C_2 < \infty.$$

(H5) h_K is a positive sequence such that

$$\lim_{n\to\infty} h_K = 0 \text{ and } \lim_{n\to\infty} \log n/n\phi_x(h_K) = 0$$

(H6) h_H is positive sequence such that

$$\lim_{n\to\infty} h_H = 0 \quad \text{with} \lim_{n\to\infty} n^\alpha h_H = \infty, \text{for some } \alpha > 0$$

Theorem 3.8 *(Ferraty et al.(2006))*
under (H1)-(H6), we obtain that

$$\sup_{y\in S_{\mathbb{R}}}|F^x(y)-\hat{F}^x(y)| = O(h_K^{b_1}) + O(h_H^{b_2}) + O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right)$$

3.2.2 Local linear estimation of conditional distribution function

We are going to adopt the model of the conditional cumulative distribution function estimator with local linear estimation method which is proposed by Demongeot et al.(2014) (see the previous chapter paragraph (1.6.2)). In what follows, we are going to give some asymptotic results that related to the behaviour of our estimator.

Assumptions and asymptotic results

(H1) $\forall r>0,\ \phi_x(r) := \phi_x(-r,r) > 0.$

(H2) The conditional cumulative distribution function F^x satisfies : there exist $b_1>0, b_2>0$, $\forall(y_1,y_2)\in S_{\mathbb{R}}\times S_{\mathbb{R}}, \forall(x_1,x_2)\in N_x\times N_x$

$$|F^{x_1}(y_1)-F^{x_2}(y_2)| \leq C(d^{b_1}(x_1,x_2)+|y_1-y_2|^{b_2}).$$

(H3) The function $\beta(.,.)$ is such that:

$$\forall z\in\mathcal{F}, C_1 d(x,z) \leq| \beta(x,z) |\leq C_2 d(x,z)$$

where $C_1>0,\ C_2>0,$ and $d(x,z)=|\delta(x,z)|$.

(H4) K is a positive, differentiable function with support $(-1,1)$.

(H5) The kernel H is a differentiable function, such that:

$$\int |t|^{b_2}H^{(1)}(t)\mathrm{d}t < \infty.$$

$(H6)$ The bandwidth h_k is such that : there exists a positive integer n_0, such that $\forall n > n_0$

$$-\frac{1}{\phi_x(h_K)}\int_{-1}^{1}\phi_x(zh_K,h_K)\frac{d}{dz}(z^2K(z))dz > C_3 > 0$$

$$\lim_{n\to\infty}\frac{\log n}{n\phi_x(h_K)} = 0;$$

and

$$h_K\int_{B(x,h_K)}\beta(u,x)dP(u) = o\left(\int_{B(x,h_K)}\beta^2(u,x)dP(u)\right)$$

where $B(x,r)$ denotes the closed-ball and $dP(x)$ is the cumulative distribution of X.

Theorem 3.9 *The rate of almost complete convergence(Demongeot et al.(2014)) under* $(H1)$*-*$(H6)$*, we obtain that:*

$$|\hat{F}^x(y) - F^x(y)| = O(h_k^{b_1} + h_H^{b_2}) + O_{a.co}\left(\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right)$$

In this part, we are going to introduce the study of Demongeot et al.(2014), in which they studied the L^2-consistency of $\hat{F}^x(y)$, to do that, they proposed for any $l \in \{0,2\}$:

$$\psi_l(x,y) = \frac{\partial^l F^x(y)}{\partial y^l} \text{ and } \Psi_l(s) = \mathbb{E}[\psi_l(X,y) - \psi_l(x,y)|\beta(x,X) = s],$$

$$B_H(x,y) = \frac{1}{2}\frac{\partial^2 F^x(y)}{\partial y^2}\int t^2 H^{(1)}(t)dt,$$

$$B_K(x,y) = \frac{1}{2}\Psi_0^2(0)\frac{\left(K(1) - \int_{-1}^{1}(u^2K(u))'\tau_x(u)du\right)}{\left(K(1) - \int_{-1}^{1}K'(u)\tau_x(u)du\right)} + o(h_K^2),$$

and

$$V_{HK}(x,y) = F^x(y)(1 - F^x(y))\left[\frac{\left(K^2(1) - \int_{-1}^{1}(K^2(u))'\tau_x(u)du\right)}{\left(K(1) - \int_{-1}^{1}(K(u))'\tau_x(u)du\right)^2}\right].$$

and they set the following hypotheses

Assumptions and asymptotic results

(M1) The hypothesis (H1) is fulfilled and there exists a function $\tau_x(.)$ such that :

$$\forall t \in (-1,1), \lim_{h_K \to 0} \frac{\phi_x(th_K, h_K)}{\phi_x(h_K)} = \tau_x(t).$$

(M2) For any $l \in \{0,2\}$, the quantities $\Psi_l^{(2)}(0)$ exist, where $\Psi_l^{(k)}$ denotes the kth order derivative of Ψ_l.

(M3) The function $\beta(.,.)$ satisfies (H3) and :

$$\sup_{u \in B(x,r)} |\beta(u,x) - \delta(u,x)| = o(r).$$

(M4) The condition (H4) holds, and the first derivative of the kernel K satisfies :

$$K^2(1) - \int_{-1}^{1} (K^2(u))' \tau_x(u) du > 0.$$

(M5) The kernel H satisfies (H5) and its first derivative $H^{(1)}$ is symmetric and such that:

$$\int t^2 H^{(1)}(t) dt < \infty$$

Theorem 3.10 *Demongeot et al.(2014) Under assumptions (M1)-(M5) and (H6), we have that*

$$\mathbb{E}\left[\hat{F}^x(y) - F^x(y)\right]^2 = B_H^2(x,y)h_H^4 + B_K^2 h_K^4 + \frac{V_{HK}(x,y)}{n\phi_x(h_K)} + o(h_H^4 + h_K^4) + o\left(\frac{1}{n\phi_x(h_K)}\right).$$

Under the previous assumptions and notations, Bouanani et al (2019) established the asymptotic normality of the local linear estimator of the conditional cumulative distribution function in iid case.

Theorem 3.11 *Bouanani et al (2019)*
Assume that : $h_K \to 0$, $h_H \to 0$. Under (M1), (H2), (M3),(M4), (H5) and (H6) we have

$$\sqrt{n\phi_x(h_K)}\left(\hat{F}^x(y) - F^x(y) - B_n(x,y)\right) \xrightarrow{D} \mathcal{N}(0, V_{HK}(x,y))$$

where

$$B_n(x,y) = \frac{\mathbb{E}[\hat{F}_N^x(y)]}{\mathbb{E}[\hat{F}_D^x]} - F^x(y),$$

$$\hat{F}_N^x(y) = \frac{1}{n\mathbb{E}(\Delta_1 K_1)}\sum_{j=1}^{n}\Delta_j K_j H_j \quad \text{and} \quad \hat{F}_D^x = \frac{1}{n\mathbb{E}(\Delta_1 K_1)}\sum_{j=1}^{n}\Delta_j K_j,$$

with

$$\Delta_j = \sum_{i=1}^{n}\beta_i^2 K_i - \left(\sum_{i=1}^{n}\beta_i K_i\right)\beta_j$$

3.3 The conditional density

To satisfy our curiosity and show more asymptotic results, this present section is organized to present the conditional density function's results.

3.3.1 Kernel conditional density function estimator and its derivatives

Through this paragraph, we will study the problem of the conditional density function estimator's behaviour; therefore, in this investigation we can be based on the results of Ferraty et al.(2006) and Laksaci (2007).
Ferraty et al.(2006) introduced the following Kernel type estimators for the successive derivatives of the conditional density :

$$\hat{f}^{x(j)}(y) = \frac{h_H^{-j-1}\sum_{i=1}^{n} K\left(h_K^{-1}d(x,X_i)\right) H^{(j+1)}\left(h_H^{-1}(y-Y_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1}d(x,X_i)\right)}.$$

Assumptions and asymptotic results

To get the rate of almost-complete convergence of the functional kernel estimator $\hat{f}^{x(j)}$, the previous model satisfies these conditions (see Ferraty et al.(2006)) :

(H1) for all $\epsilon > 0, \mathbb{P}(X \in B(x,\epsilon)) = \phi_x(\epsilon) > 0$.

(H2) $\forall(x_1, x_2) \in N_x^2, \forall(y_1, y_2) \in S_{\mathbb{R}}^2$,

$$|f^{x_1(j)}(y_1) - f^{x_2(j)}(y_2)| \leq C_x(d(x_1, x_2)^{b_1} + |y_1 - y_2|^{b_2})$$

(H3) $\begin{cases} \forall(y_1, y_2) \in \mathbb{R}^2, |H(y_1) - H(y_2)| \leq C|y_1 - y_2|, \\ \int |t|^{b_2} H^{(1)}(t)dt < \infty. \end{cases}$

(H4) The kernel K is a function with support $(0, 1)$ such that

$$0 < C_1 < K(t) < C_2 < \infty.$$

(H5) h_K is a positive sequence such that

$$\lim_{n\to\infty} h_K = 0 \text{ and } \lim_{n\to\infty} \frac{\log n}{n h_H^{2j+1} \phi_x(h_K)} = 0.$$

(H6) h_H is positive sequence such that
$\lim_{n\to\infty} h_H = 0$ with $\lim_{n\to\infty} n^\alpha h_H = \infty$, for some $\alpha > 0$

(H7) $\begin{cases} \forall(y_1, y_2) \in \mathbb{R}^2, |H^{(j+1)}(y_1) - H^{(j+1)}(y_2)| \leq C|y_1 - y_2|, \\ \exists \nu > 0, \forall j' \leq j+1, \lim_{y\to\infty} |y|^{1+\nu} |H^{(j+1)}(y)| = 0, \\ H^{(j+1)} \text{is bounded} \end{cases}$

Theorem 3.12 *the rate of almost-complete convergence (Ferraty et al.(2006)) Under the hypotheses (H1)-(H7), we get*

$$\sup_{y\in S_{\mathbb{R}}} |\hat{f}^{x(j)}(y) - f^{x(j)}(y)| = O(h_K^{b_1}) + O(h_H^{b_2}) + O\left(\sqrt{\frac{\log n}{n h_H^{2j+1} \phi_x(h_K)}}\right), \; a.co.$$

And for more statistic results, Laksaci (2007) stated the mean square convergence from next additional hypotheses

Assumptions and asymptotic results

(M1) $\forall r > 0$, the random variable $Z = r^{-1}(x - X)$ is absolutely continuous with respect to the measure μ. Its density $g(r, x, v)$ is strictly positive on $B(0, 1)$ such that :

$$g(r, x, v) = \phi(r)h(x, v) + o(\phi(r)) \text{ for all } v \in B(0, 1),$$

where : ϕ is an increasing function within $\mathbb{R}^+$, $h : \mathcal{F} \times \mathcal{F} \to \mathbb{R}^+$ and

$$0 < \int_{B(0,1)} h(x,v)d\mu(v) < \infty.$$

(M2) The kernel H is a positive, bounded, integrable, symmetric and :

$$\int H(t)dt = 1 \text{ and } \int t^2 H(t)dt < \infty.$$

(M3) $\lim_{n\to\infty} h_K = 0$, $\lim_{n\to\infty} h_H = 0$ and $\lim_{n\to\infty} nh_H\phi(h_K) = \infty$.

Theorem 3.13 *mean square convergence (Laksaci (2007))*
if $f^x \in C_B^2(\mathcal{F} \times \mathbb{R})$ *and under hypotheses (M1)-(M3) and (H4), we have that*

$$\begin{aligned} \mathbb{E}\left[\hat{f}^x(y) - f^x(y)\right]^2 &= B_H^2(x,y)h_H^4 + B_K^2(x,y)h_K^4 + \tfrac{V_{HK}(x,y)}{nh_H\phi_x(h_K)} \\ &+o(h_H^4 + h_K^2) + o\left(\tfrac{1}{nh_H\phi_x(h_K)}\right) \end{aligned}$$

with

$$B_H(x,y) = \tfrac{1}{2}\tfrac{\partial^2 f^x(y)}{\partial y^2}\int t^2 H(t)dt, \quad B_K(x,y) = \frac{\int_{B(0,1)} K(||v||)D_x f^x(y)[v]h(x,v)d\mu(v)}{\int_{B(0,1)} K(||v||)h(x,v)d\mu(v)}$$

$$V_{HK}(x,y) = (f^x(y))\left(\int_{B(0,1)} K(||v||)h(x,v)d\mu(v)\right)^{-2}\left(\int_{B(0,1)} K^2(||v||)h(x,v)d\mu(v)\right)\int H^2(t)dt$$

where D_x is the derivative with respect to x.

3.3.2 Local linear estimation of conditional density

According to Demongeot et al.(2013), we will present both the pointwise and the uniform almost-complete consistencies with convergence rates of the conditional density estimator.

Assumptions and asymptotic results

At the beginning of this part, we try to introduce the rate of almost-complete convergence which established by Demongeot et al.(2013) under the following constraints :

(H1) $\forall r > 0,\ \phi_x(r) := \phi_x(-r, r) > 0.$

(H2) The conditional distribution function f^x is such that : there exist $b_1 > 0, b_2 > 0, \forall (y_1, y_2) \in S_{\mathbb{R}} \times S_{\mathbb{R}}, \forall (x_1, x_2) \in N_x \times N_x$

$$|f^{x_1}(y_1) - f^{x_2}(y_2)| \leq C_x(d^{b_1}(x_1, x_2) + |y_1 - y_2|^{b_2}),$$

C_x is a positive constant depending on x.

(H3) The function $\beta(.,.)$ is such that :

$$\forall z \in \mathcal{F}, C_1 d(x, z) \leq |\beta(x, z)| \leq C_2 d(x, z),$$

where $C_1 > 0 \;\; C_2 > 0,$ and $d(x, z) = |\delta(x, z)|.$

(H4) K is a positive, differentiable function with support $(-1, 1)$.

(H5) The kernel H is a positive, bounded and Lipschitzian continuous function, such that :

$$\int |t|^{b_2} H(t)\mathrm{d}t < \infty \quad \text{and} \quad \int H^2(t)\mathrm{d}t < \infty.$$

(H6) The bandwidth h_k satisfies : $\exists n_0 \in \mathbb{N}$, such that :

$$\forall n > n_0, -\frac{1}{\phi_x(h_K)} \int_{-1}^{1} \phi_x(zh_K, h_K) \frac{d}{dz}(z^2 K(z))dz > C_3 > 0,$$

and

$$h_K \int_{B(x,h_K)} \beta(u, x)dP(u) = o\left(\int_{B(x,h_K)} \beta^2(u, x)dP(u) \right),$$

where $B(x, r)$ denotes the closed-ball and $dP(x)$ is the cumulative distribution of X.

(H7) The bandwidth h_H is such that : for some $\gamma > 0$

$$\lim_{n\to\infty} n^{\gamma} h_H = \infty \quad \text{and} \quad \lim_{n\to\infty} \frac{\ln n}{n h_H \phi_x(h_K)} = 0.$$

Theorem 3.14 *rate of pointwise almost complete convergence(Demongeot et al.(2013)) Under (H1)-(H7), we obtain*

$$\sup_{y\in S_{\mathbb{R}}}|\hat{f}^x(y)-f^x(y)| = O(h_k^{b_1}+h_H^{b_2})+O\left(\sqrt{\frac{\ln n}{nh_H\phi_x(h_K)}}\right) \quad a.co.$$

As the uniform almost-complete convergence require some additional tools and topological conditions,hence, Demongeot et al.(2013) exhibited the next conditions
Assumptions and asymptotic results

(H1′) There exists a differentiable function $\phi(.)$, such that:
$\forall x \in S_{\mathcal{F}}, 0 < C\phi(h) \leq \phi_x(h) \leq C^{'}\phi(h) < \infty$ and $\exists \eta_0 > 0, \forall \eta < \eta_0, \phi^{(1)}(\eta) < C$.

(H2′) The conditional density function f^x is such that: $\forall(y_1, y_2) \in S_{\mathbb{R}}\times S_{\mathbb{R}}, \forall(x_1, x_2) \in S_{\mathcal{F}} \times S_{\mathcal{F}}$

$$|f^{x_1}(y_1) - f^{x_2}(y_2)| \leq C(d^{b_1}(x_1, x_2) + |y_1 - y_2|^{b_2}),$$

C is strictly positive constant.

(H3′) Under condition (H3) and, for some strictly positive constant $C^{'}$, the function $\beta(.,.)$ satisfies the following Lipschitz's condition:

$$\forall(x_1, x_2) \in S_{\mathcal{F}} \times S_{\mathcal{F}} : |\beta(x_1, x^{'}) - \beta(x_2, x^{'})| \leq C^{'}d(x_1, x_2).$$

(H4′) The condition (H4) is satisfied and, for some strictly positive constant C, the kernel K satisfies the following Lipschitz's condition:

$$|K(x) - K(y)| \leq C||x| - |y||.$$

(H5′) For $r_n = O\left(\frac{\ln n}{n}\right)$, and for some $\gamma \in (0, 1), \lim_{n\to\infty} n^{\gamma}h_H = \infty$ the sequence d_n satisfies :

$$\frac{(\ln n)^2}{n^{1-\gamma}\phi(h_k)} < \ln d_n < \frac{n^{1-\gamma}\phi(h_k)}{\ln n},$$

and:

$$\sum_{n=1}^{\infty} n^{\frac{(3\gamma+1)}{2}} d_n^{1-\beta} < \infty, \text{ for some } \beta > 1.$$

Theorem 3.15 *rate of uniform almost complete convergence (Demongeot et al.(2013)) Under assumptions (H1′)-(H5′), (H5) and (H6), we have that*

$$\sup_{x\in S_{\mathcal{F}}}\sup_{y\in S_{\mathbb{R}}}|\hat{f}^x(y)-f^x(y)| = O(h_k^{b_1}+h_H^{b_2}) + O_{a.co}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right).$$

On the other side, Rachdi et al.(2014) set for some $l\in\{0,2\}$

$$\psi_l(.,y) = \frac{\partial^l f^x(y)}{\partial y^l}$$
$$\text{and}\quad \Psi_l(s) = \mathbb{E}[\psi_l(X)-\psi_l(x)|\delta(x,X)=s].$$

Before giving the mean square convergence, we take care to our model by giving the following hypotheses (see Rachdi et al.(2014))

Assumptions and asymptotic results

(M1) $\forall r>0, \phi_x(r) := \phi_x(-r,r) > 0$ and there exists a function$\tau_x(.)$such that

$$\forall t\in(-1,1), \lim_{h\to 0}\frac{\phi_x(-h,th)}{\phi_x(h)} = \tau_x(t).$$

(M2) For $l\in\{0,2\}$, the quantities $\Psi_l'(0)$ and $\Psi_l''(0)$ exist.

(M3) The function $\beta(.,.)$ satisfies (H3), the second part of (H6) and the following condition :

$$\sup_{u\in B(x,r)}|\beta(u,x)-\delta(u,x)| = o(r).$$

(M4) K is a positive, differentiable function which is supported within $(-1,1)$.Its derivative K' satisfies $K'(t)<0$, for $-1\le t<1$, and $K(1)>0$.

(M5) The kernel H is a positive, integrable, bounded, symmetric and :

$$\int H(t)dt = 1 \text{ and } \int t^2 H(t)dt < \infty$$

(M6) The bandwidths h_K and h_H satisfy: $\lim_{n\to\infty} h_K = 0,\quad \lim_{n\to\infty} h_H = 0.$
Moreover, $\lim_{n\to\infty} n h_H \phi(h_K) = \infty$.

Theorem 3.16 *mean square convergence (Rachdi et al.(2014))*
Under hypotheses (M1)-(M6), we have that

$$\begin{aligned} \mathbb{E}\left[\hat{f}^x(y) - f^x(y)\right]^2 &= B_H^2(x,y)h_H^4 + B_K^2(x,y)h_K^4 \\ &+ \tfrac{V_{HK}(x,y)}{nh_H\phi_x(h_K)} + o(h_H^4 + h_K^2) + o\left(\tfrac{1}{nh_H\phi_x(h_K)}\right) \end{aligned}$$

with

$$B_H(x,y) = \frac{1}{2}\psi_2(x,y)\int t^2 H(t)dt$$

$$B_K(x,y) = \frac{1}{2}\Psi_0''(0)\left[\frac{K(1) - \int_{-1}^{1}(u^2K(u))'\tau_x(u)du}{K(1) - \int_{-1}^{1}(K(u))'\tau_x(u)du}\right]$$

$$V_{HK}(x,y) = f^x(y)\left(\int H^2(t)dt\right)\left[\frac{\left(K^2(1) - \int_{-1}^{1}(K^2(u))'\tau_x(u)du\right)}{\left(K(1) - \int_{-1}^{1}(K(u))'\tau_x(u)du\right)^2}\right]$$

References

BAÍLLO, A. and GRANÉ, A. (2009). Functional Local Linear Regression with Functional Predictor and Scalar Response. *Journal of Multivariate Analysis, **100**, 102-111.*

BARRIENTOS-MARIN, J., FERRATY, F. *and* VIEU, P. *(2010). Locally Modelled Regression and Functional Data.* J. of Nonparametric Statistics ***22**, 617-632.*

BOUANANI, O.*,* LAKSACI, A.*,* RACHDI, M. *and* RAHMANI, S. *(2019). Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics.* Behaviormetrika, **46**, 199-233.

DABO-NIANG, S. and RHOMARI, N. (2003). Kernel regression estimation when the regressor takes values in metric space. *C. R. Acad. Sci. Paris*, **336**, 75-80.

DEMONGEOT, J., LAKSACI, A., MADANI, F. and RACHDI, M. (2013). Functional data : local linear estimation of the conditional density and its application.*Statistics*, **47**, 26-44.

DEMONGEOT, J., LAKSACI, A. RACHDI, M. and RAHMANI, S. (2014). On the local Modalization of the conditional distribution for functional data. *Sankhya A*, **76** (2), 328-355.

FERRATY, F.,LAKSACI, A. andVIEU, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models. *Statistical Inference for Stochastic Processes*, **9**, 47-76.

FERRATY, F. and VIEU, P. (2006). *Nonparametric functional data analysis. Theory and Practice.* Springer Series in Statistics. New York.

LAKSACI, A. (2007). Erreur quadratique de l'estimateur à noyau de la densité conditionnelle à variable explicative fonctionnelle. *C. R. Acad. Sci. Paris,***345**, 171-175.

Rachdi, M., Laksaci, A.,Demongeot, J., Abdali, A. and Madani, F. (2014). Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data. Computational Statistics and Data Analysis, **73**, 53-68.

Zhou, Z.Lin, Z.(2015). Asymptotic normality of locally modelled regression estimator for functional data. Journal of Nonparametric Statistics, **28**, 116-131.

Chapter 4

Local linear estimation of c.d.f in the functional data: Uniform consistency with convergence rates.

This chapter is submitted in South African Statistical Journal

Local linear estimation of c.d.f in the functional data: Uniform consistency with convergence rates.

Chaima Hebchi[1], Abdelhak CHOUAF[2]

[1,2]Laboratoire de Statistique et Processus Stochastiques, (LSPS) Université Djillali liabès, BP 89, Sidi bel Abbès 22000, Algeria

E-mail: [1]chaimahabchi@yahoo.fr

Abstract

This paper aims to state the rate of uniform almost complete convergence linked with nonparametric estimation of conditional cumulative distribution function and we take the conditional mode as an application

Keywords: Functional data, Local linear estimator,Conditional cumulative, Conditional mode, Nonparametric regression, Small balls probability.

2000 Mathematics Subject Classification: 62G05, 62G07, 62G20

4.1 Introduction

Conditional estimation is an important field and useful in all domains of statistics, such as time series, survival analysis and growth charts among others, for more information see : Stone (1977) and Koenker (2000, 2005). There exists an extensive literature and various nonparametric approaches in conditional estimation for independent samples and dependent non-functional or functional observations. Among a lot of papers that are dealing with conditional estimation in finite dimension, one can refer for example to key works of Stute (1986), Portnoy (1991), Koul and Mukherjee (1994), Honda (2000), Yu et al. (2003)

Noting that, these questions of the modelization statistic of functional data has known a growing interest among theoretical and applied statisticians (see Bosq (2000), Ramsay and Silverman (2002-2005) also Ferraty and Vieu(2006)). In this context, the conditional cumulative distribution function and its derivative (the conditional density function) have a great importance in many applications such as reliability and survival analysis. Moreover, they provide information about the relationship between X and Y furthermore, they lead to some prediction method, such as the conditional mode, the conditional median or the conditional quantiles (for more details see Ferraty et al. (2010) for a list of references). As is well known, the local polynomial smoothing has various advantages over the kernel method (see Fan and Gijbels (1996) for an extensive discussion on the comparison between both methods, in the multivariate case). In the nonfunctional case, the local polynomial fitting has been the subject of considerable study. Besides, there are valuable references on this topic such as Fan (1992), Fan and Gijbels (1996) and Fan and Yao (2003). However, only few results are available for local linear modelling in functional statistics. Indeed, the first result in this topic was obtained by Baíllo and Grané(2009). They studied the local linear estimator of the regression function when the explanatory variable takes values in a Hilbert space. However, The general case where regressors are not Hilbertian has been considered by Barrientos-Marin et al. (2010). Recall that in the i.i.d setting, Barrientos-Marin et al.(2010) introduced the local linear estimator of the regression operator of a scalar response Y on an explanatory functional variable X, this method had sevral adventages like making the estimator computation easy and fast while keeping good predictive performance (see Xiong et al.(2017)). In this pioneering work, the authors obtained the almost complete convergence (with rate) of the proposed estimate. We return to Boj et al. (2010) for an other alternative version for the functional local linear modelling. More recently, Demongeot et al. (2011) consider the local polynomial modelling of the conditional density function when the explanatory variable is functional and the quadratic error of this estimator has been treated by Rachdi et al.(2014). Thereafter, the almost-complete convergence with rates of the local linear estimator of the conditional cumultative distribution is stated by Demongeot et al.(2014).

In the iid setting, our work focuses on the local linear estimation of the conditional cumultative distribution for functional data. In section 2, we started by clarifing our model and under some assumptions the main asymptotic results are stated in section3. In section 4, we will exploit these results to the conditional mode estimation.

4.2 Model and assumptions

At this stage, we observe n pairs (X_i, Y_i) for $i = 1, ..., n$ identically distributed as (X, Y), this last is valude in $\mathcal{F} \times \mathbb{R}$, where $\mathcal{F}$ is a semi-metric space equipped with a semi-metric d. We suppose for $x \in \mathcal{F}$ that there exists a regular version of conditional probability of Y given $X = x$, which is absolutely continous with respect to the Lebesgue measure on $\mathbb{R}$. the functional local linear estimator $\hat{F}^x(.)$ is defined by :

$$\hat{F}^x(y) = \frac{\sum_{i,j=1}^{n} W_{i,j}(x) H(h_H^{-1}(y - Y_i))}{\sum_{i,j=1}^{n} W_{i,j}(x)} \tag{4.1}$$

where

$$W_{i,j}(x) = \beta(X_i, x)(\beta(X_i, x) - \beta(X_j, x))K(h_K^{-1}\delta(x, X_i))K(h_K^{-1}\delta(x, X_j))$$

$\beta(.,.)$ and $\delta(.,.)$ are two functions defined from $\mathcal{F} \times \mathcal{F}$ to $\mathbb{R}$, such that: $\forall \xi \in \mathcal{F}, \beta(\xi, \xi) = 0$, and $d(.,.) = |\delta(.,.)|$. K and H are Kernels, $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) is chosen as a sequence of positive real numbers and each of them converges to 0 when $n \to \infty$.
with the convention $0/0 = 0$. We aim to state the uniform almost complete convergence of $\hat{F}$ on some subset $S_{\mathcal{F}}$ of $\mathcal{F}$, where:

$$S_{\mathcal{F}} \subset \cup_{k=1}^{d_n} B(x_k, r_n)$$

where $x_k \in \mathcal{F}$ and r_n (resp d_n) is a sequence of positive real numbers, $\phi_x(r_1, r_2) = \mathbb{P}(r_2 \leq \delta(X, x) \leq r_1)$. We assume that our nonparametric model satisfies the following conditions:

$(H1)$ There exists a differentiable function $\phi(.)$, such that:
$\forall x \in S_{\mathcal{F}}, 0 < C\phi(h) \leq \phi_x(h) \leq C'\phi(h) < \infty$ and $\exists \eta_0 > 0, \forall \eta < \eta_0$,
$\phi^{(1)}(\eta) < C$,

$(H2)$ The conditional distribution function F^x is such that:
there exists $b_1 > 0, b_2 > 0, \forall (y_1, y_2) \in S_{\mathbb{R}} \times S_{\mathbb{R}}, \forall (x_1, x_2) \in S_{\mathcal{F}} \times S_{\mathcal{F}}$

$$|F^{x_1}(y_1) - F^{x_2}(y_2)| \leq C(d^{b_1}(x_1, x_2) + |y_1 - y_2|^{b_2})$$

$(H3)$ The function $\beta(.,.)$ is such that:

$$\forall x^{'} \in \mathcal{F}, C_1 d(x,x^{'}) \leq\mid \beta(x,x^{'}) \mid\leq C_2 d(x,x^{'}) \text{ where } C_1 > 0 \text{ and } C_2 > 0$$

and, for some strictly positive constant $C^{'}$, the following Lipschitz's condition:

$$\forall (x_1,x_2) \in S_{\mathcal{F}} \times S_{\mathcal{F}} : |\beta(x_1,x^{'}) - \beta(x_2,x^{'})| \leq C^{'} d(x_1,x_2)$$

$(H4)$ K is a positive, differentiable function with support $[-1,1]$ and, for some strictly positive constant C, the following Lipschitz's condition:

$$\mid K(x) - K(y) \mid\leq C \mid\mid x \mid - \mid y \mid\mid$$

$(H5)$ The kernel H is a differentiable, positive, bounded and Lipschizian function, such that: H is of classe C^2, of compact support and satisfies:

$$\int |t|^{b_2} H^{(1)}(t)\mathrm{dt} < \infty$$

$(H6)$ The bandwidth h_k satisfies: there exists an integer n_0, such that:
$\forall n > n_0, -\frac{1}{\Phi_x(h_K)} \int_{-1}^{1} \phi_x(zh_K,h_K)\frac{d}{dz}(z^2K(z))dz > C_3 > 0$
and
$h_K \int_{B(x,h_K)} \beta(u,x)dP(u) = o\left(\int_{B(x,h_K)} \beta^2(u,x)dP(u)\right)$
where $B(x,r)$ denotes the clossed-ball and $dP(x)$ is the cumultative distribution of X

$(H7)$ For $r_n = O\left(\frac{\ln n}{n}\right)$, and for some $\gamma \in (0,1)$, for n large enough the sequence d_n satisfies :

$$\frac{(\ln n)^2}{n^{1-\gamma} h_H^2 \phi(h_k)} < \ln d_n < \frac{n^{1-\gamma} h_H^2 \phi(h_k)}{\ln n}$$

and:

$$\sum_{n=1}^{\infty} n^{\frac{(3\gamma+1)}{2}} d_n^{1-\beta} < \infty, \text{ for some } \beta > 1$$

Obviously, these conditions are commonly used in many studies of the local linear method for functional data.

4.3 Asymptotic results

Before given the asymptotic result,we introduce the following notations $\hat{F}_N^x(y) = \dfrac{1}{n(n-1)EW_{12}} \sum_{i,j=1}^{n} W_{i,j} H_i(h_H^{-1}(y-Y_i)$ and $\hat{f}_D^x = \dfrac{1}{n(n-1)EW_{12}} \sum_{i,j=1}^{n} W_{i,j}$

Theorem 4.1 *under* $(H1)$-$(H7)$, *we obtain that:*

$$\sup_{x\in S_{\mathcal{F}}} \sup_{y\in S_{\mathbb{R}}} |\hat{F}^x(y) - F^x(y)| = O(h_k^{b_1} + h_H^{b_2}) + O_{a.co}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right)$$

Remark that, the theorem's proof can be deduced directly from the following decomposition :

$$\begin{aligned} \hat{F}^x(y) - F^x(y) &= \tfrac{1}{\hat{f}_D^x}\left\{\left(\hat{F}_N^x(y) - \mathbb{E}[\hat{F}_N^x(y)]\right) - \left(F^x(y) - \mathbb{E}[\hat{F}_N^x(y)]\right)\right\} \\ &+ \tfrac{\hat{F}^x(y)}{\hat{f}_D^x}\left(1 - \hat{f}_D^x\right) \end{aligned} \tag{4.2}$$

in addition the following lemmas (for which the proofs are given in the Appendix) lead us to get Theorem 4.1.

Lemma 4.1 *Demongeot et al. (2013) under assumptions* $(H1)$, $(H3)$, $(H4)$, $(H5)$ *and* $(H6)$, *we obtain that :*

$$\sup_{x\in S_{\mathcal{F}}} |\hat{f}_D^x - 1| = O_{a.co}\left(\sqrt{\frac{\ln d_n}{n\phi(h_K)}}\right)$$

Corollary 4.1 *Demongeot et al. (2013) Under the assumptions of Lemma 6.1, we have that :*

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{x\in S_{\mathcal{F}}} \hat{f}_D^x < \frac{1}{2}\right) < \infty$$

Lemma 4.2 *Demongeot et al. (2014)Under assumptions* $(H1)$, $(H2)$, $(H4)$ *and* $(H5)$ *we obtain that :*

$$\sup_{x\in S_{\mathcal{F}}} \sup_{y\in S_{\mathbb{R}}} |F^x(y) - \mathbb{E}[\hat{F}_N^x(y)]| = O(h_K^{b_1}) + O(h_H^{b_2})$$

Lemma 4.3 *under the hypotheses* $(H1)-(H7)$ *we obtain that:*

$$\sup_{x\in S_F}\sup_{y\in S_{\mathbb{R}}}|\hat{F}_N^x(y)-\mathbb{E}\hat{F}_N^x(y)| = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right)$$

4.4 Application: Conditional mode estimation

The purpose of this section is to state the almost complete convergence of the conditional mode of Y given $X=x$, denoted by $\theta(x)$. For this aim, we will need the following assumptions :

$(H2^{'})$ for all $(x_1,x_2)\in S_{\mathcal{F}}\times S_{\mathcal{F}}$ and $(y_1,y_2)\in S_{\mathbb{R}}\times S_{\mathbb{R}}$, we have :

$$|f^{x_1}(y_1)-f^{x_2}(y_2)|\le C(d^{b_1}(x_1,x_2)+|y_1-y_2|^{b_2})$$

for some positive constants b_1 and $b_2>0$, where C is a positive constant.

$(H5^{'})$ H is a positive, bounded, Lipschitizian continuous function, such that :

$$\int |t|^{b_2}H(t)dt<\infty \text{ and } \int H^2(t)dt<\infty.$$

$(H8)$ The sequence d_n satisfies $(H7)$ and, $\lim_{n\to+\infty} n^{\gamma}h_H=\infty$

$(H9)$ we assume that ; $f^x\in C^0_{\mathcal{F}\times\mathbb{R}}\cap\mathcal{S}^x_{dens}$ (the continuity-type functional nonparametric model cf. Ferraty and Vieu (2006)), where

$$C^0_{\mathcal{F}\times\mathbb{R}}=\left\{\begin{array}{c} f:\mathcal{F}\times\mathbb{R}\to\mathbb{R},\forall x^{'}\in S_{\mathcal{F}} \\ \lim_{d(x,x^{'})\to 0} f(x^{'},y)=f(x,y), \\ \forall y^{'}\in\mathbb{R},\ \lim_{|y-y^{'}|\to 0} f(x,y^{'})=f(x,y)\end{array}\right\}$$

and

$$\mathcal{S}^x_{dens}=\left\{\begin{array}{l} f:\mathcal{F}\times\mathbb{R}\to\mathbb{R}; \\ \exists\zeta>0,\exists! y_0\in S_{\mathbb{R}}, f(x,.)\text{is strictly increasing on} \\ (\theta_f-\zeta,\theta_f)\text{and strictly decreasing on}(\theta_f,\theta_f+\zeta)\end{array}\right\}$$

$(H10)$ $\forall \epsilon_0 > 0, \exists \eta > 0, \forall r : S \to S_{\mathbb{R}}$, we have that :

$$\sup_{x\in S_{\mathcal{F}}} |\theta_f(x) - r(x)| \geq \epsilon_0 \Rightarrow \sup_{x\in S_{\mathcal{F}}} |f^x(r(x)) - f^x(\theta_f(x))| \geq \eta$$

$(H11)$ there exists some integer $j > 1$ such that $\forall x \in S_{\mathcal{F}}$, the function f^x is j-times continuously differentiable on interior $(S_{\mathbb{R}})$ with respect to y, and :

$$\begin{cases} f^{x(l)}(\theta_f(x)) = 0, \text{ if } 1 \leq l < j \\ \text{and} f^{x(j)}(.) \text{is uniformly continuous on } S_{\mathbb{R}} \\ \text{such that} |f^{x(j)}(\theta_f(x))| > C > 0 \end{cases}$$

where $f^{x(j)}$ denotes the jth order derivative of the conditional density f^x.

The estimator of $\theta_f(x)$ is the random variable $\hat{\theta_f}(x)$ which defined by

$$\hat{\theta_f}(x) = \arg\min_{y\in S_{\mathbb{R}}} \hat{f}^x(y)$$

and from Theorem 3.4.1 in Demongeot et al. (2013), we can get the following corollary

Corollary 4.2 *Under the hypotheses* $(H1)$,$(H2')$,$(H3)$, $(H4)$, $(H5')$ *and* $(H6)-(H11)$, *we get*

$$\sup_{x\in S_{\mathcal{F}}} |\hat{\theta_f}(x) - \theta_f(x)|^j = O(h_K^{b_1} + h_K^{b_2}) + O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma\phi(h_K)}}}\right)$$

4.5 Appendix

In what follows, when no confusion is possible, we put for any $x \in \mathcal{F}$, and for all $i = 1, ..., n$:

$$K_i(x) = K(h^{-1}\delta(x, X_i)), \beta_i(x) = \beta(X_i, x) \text{ and } H_i(y) = H(h_H^{-1}(y - Y_i))$$

Proof of Lemma 4.3. The proof of this lemma follows the same steps as in [Demongeot et al. (2013), lemma 4.2] , where $S_2(x)$, $S_3(x)$ and $S_4(x)$ are replaced by :

$$\begin{cases} M_2^x(y) &= \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}\dfrac{K_j(x)H_j(y)}{\phi_x(h_K)} \\ M_3^x(y) &= \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}\dfrac{K_j(x)\beta_j(x)H_j(y)}{h_K\phi_x(h_K)} \\ M_4^x(y) &= \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}\dfrac{K_j(x)\beta_j^2(x)H_j(y)}{h_K^2\phi_x(h_K)} \end{cases}$$

by using the compactness property of $S_{\mathbb{R}}$, we can write that: there exists a sequence of real numbers $(t_k)_{k=1,\ldots,s_n}$, such that: $S_{\mathbb{R}} \subset \cup_{k=1}^{s_n}(t_k - l_n, t_k + l_n)$ where: $l_n = n^{-\frac{3}{2}\gamma-1/2}$ and $s_n = O(l_n^{-1})$.
Taking: $t_y = \arg\min_{t\in\{t_1,\ldots,t_{s_n}\}}|y-t|$ and by $j(x) = \arg\min_{j\in\{1,2,\ldots,d_n\}}|\delta(x,x_k)|$ we consider the following decomposition :

$$\begin{aligned} |M_i^x(y) - \mathbb{E}[M_i^x(y)]| \leq & \underbrace{\sup_{x\in S_{\mathcal{F}}}\sup_{y\in S_{\mathbb{R}}}|M_i^x(y) - M_i^{x_{j(x)}}(y)|}_{A_1} \\ & + \underbrace{\sup_{x\in S_{\mathcal{F}}}\sup_{y\in S_{\mathbb{R}}}|M_i^{x_{j(x)}}(y) - M_i^{x_{j(x)}}(t_y)|}_{A_2} \\ & + \underbrace{\sup_{x\in S_{\mathcal{F}}}\sup_{y\in S_{\mathbb{R}}}|M_i^{x_{j(x)}}(t_y) - \mathbb{E}[M_i^{x_{j(x)}}(t_y)]|}_{A_3} \\ & + \underbrace{\sup_{x\in S_{\mathcal{F}}}\sup_{y\in S_{\mathbb{R}}}|\mathbb{E}[M_i^{x_{j(x)}}(t_y)] - \mathbb{E}[M_i^{x_{j(x)}}(y)]|}_{A_4} \\ & + \underbrace{\sup_{x\in S_{\mathcal{F}}}\sup_{y\in S_{\mathbb{R}}}|\mathbb{E}[M_i^{x_{j(x)}}(y)] - \mathbb{E}[M_i^x(y)]|}_{A_5} \end{aligned}$$

Similarly to the study of the term F_1 in [Demongeot et al. (2013), Lemmma 4.2], we obtain:

$$A_1 = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right) \text{ and } A_5 = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right) \tag{4.3}$$

To treat the term A_2, we use the Lipschitz's condition on the kernel H to show :

$$|M_i^{x_{j(x)}}(y) - M_i^{x_{j(x)}}(t_y)|$$

$$\leq C\frac{1}{nh_K^l\phi(h_K)}\sum_{i=1}^{n}K_i(x_{j(x)})\beta_i^l(x_{j(x)})|H_i(y) - H_i(t_y)| \leq \frac{l_n}{h_H}S_i(x_{j(x)})$$

where $S_i(.)$ for $i = 2, 3, 4$ are treated in [Demongeot et al. (2013), proof of Lemma 4.2]. Thus by using the facts that: $l_n = n^{-\frac{3}{2}\gamma-1/2}$ and $(H7)$, we obtain:

$$A_2 = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right) \text{ and } A_4 = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right) \tag{4.4}$$

Finally, for the term A_3, we have for all $\eta > 0$ that:

$$\mathbb{P}\left(A_3 > \eta\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right)$$

$$= \mathbb{P}\left(\max_{j\in\{1,2,\ldots,s_n\}}\max_{k\in\{1,2,\ldots,d_n\}} |M_i^{x_k}(t_j) - \mathbb{E}[M_i^{x_k}(t_j)]| > \eta\sqrt{\frac{\ln d_n}{n\phi(h_K)}}\right)$$

$$\leq s_n d_n \max_{j\in\{1,2,\ldots,s_n\}}\max_{k\in\{1,2,\ldots,d_n\}} \mathbb{P}\left(|M_i^{x_k}(t_j) - \mathbb{E}[M_i^{x_k}(t_j)]| > \eta\sqrt{\frac{\ln d_n}{n\phi(h_K)}}\right)$$

To do this last probability, we use the classical Bernstein's inequality such that, we put: for $l = 0, 1$,

$$Z_i^l = \frac{1}{h_K^l\phi(h_K)}\left(K_i(x_k)H_i(t_j)\beta_i^l(x_k) - \mathbb{E}[K_i(x_k)H_i(t_j)\beta_i^l(x_k)]\right)$$

By the assumption $(H3)$, we have that $\frac{1}{h_K^l}(K_i\beta_i^l) < C$ and since $H < 1$ then, we can write:

$$|Z_i^l| \leq \frac{C}{\phi_x(h_K)} \text{ and } \mathbb{E}|Z_i^l| \leq \frac{C'}{\phi_x(h_K)}$$

So, the use of the classical Bernstein's inequality (cf. Uspensky (1937), Page 205) allows us to write for all $\eta \in (0, C'/C)$:

$$\mathbb{P}\left(|\sum_{i=1}^{n} Z_i^l| \geq \eta\sqrt{\frac{\ln\ d_n}{n\phi_x(h_K)}}\right) \leq C' d_n^{-C\eta^2}$$

Therefore, the last inequality allows to get:

$$\forall j \leq s_n, \mathbb{P}\left(|M_i^{x_k}(t_j) - \mathbb{E}[M_i^{x_k}(t_j)]| > \eta\sqrt{\frac{\ln d_n}{n\phi(h_K)}}\right) \leq 2\exp\{-C\eta^2 \ln d_n\}$$

since: $s_n = O(n^{\frac{3}{2}\gamma+1/2})$, $\frac{\ln d_n}{n\phi(h_K)} > \frac{1}{n\phi(h_K)}$ and by choosing $C\eta^2 = \beta$, one gets:

$$s_n d_n \max_{j\in\{1,2,\ldots,s_n\}}\max_{k\in\{1,2,\ldots,d_n\}} \mathbb{P}\left(|M_i^{x_k}(t_j) - \mathbb{E}[M_i^{x_k}(t_j)]| > \eta\sqrt{\frac{\ln d_n}{n\phi(h_K)}}\right) \leq C'' s_n d_n^{1-\beta}$$

By using the second part of condition $(H7)$, we obtain

$$A_3 = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n^{1-\gamma}\phi(h_K)}}\right) \tag{4.5}$$

Thus, Lemma 4.3's result can be deduced from (6.2), (6.3), (6.4) ■

Proof of Corollary 4.2. With a Taylor development of $F^{x(1)}(\hat{\theta_f}(x))$ around $\theta_f(x)$, we get :

$$\begin{aligned} F^{x(1)}(\hat{\theta_f}(x)) =& F^{x(1)}(\theta_f(x)) + \sum_{i=1}^{j-1}\frac{1}{i!}(\hat{\theta_f}(x) - \theta_f(x))^i F^{x(i+1)}(\theta_f(x)) \\ &+ \frac{1}{j!}(\hat{\theta_f}(x) - \theta_f(x))^j F^{x(j+1)}(\theta_f^{'}(x)) \end{aligned}$$

which implises that :

$$\begin{aligned} f^x(\hat{\theta_f}(x)) =& f^x(\theta_f(x)) + \sum_{i=1}^{j-1}\frac{1}{i!}(\hat{\theta_f}(x) - \theta_f(x))^i f^{x(i)}(\theta_f(x)) \\ &+ \frac{1}{j!}(\hat{\theta_f}(x) - \theta_f(x))^j f^{x(j)}(\theta_f^{'}(x)) \end{aligned}$$

because of $(H11)$, we have :

$$f^x(\hat{\theta_f}(x)) = f^x(\theta_f(x)) + \frac{1}{j!}(\hat{\theta_f}(x) - \theta_f(x))^j f^{x(j)}(\theta_f^{'}(x))$$

where $\theta_f^{'}(x)$ is lying between $\theta_f(x)$ and $\hat{\theta_f}(x)$. Due to $(H9)$, $(H10)$ and $(H11)$ and by the same steps as in [Demongeot et al. (2013),corollary (5.1)] we get the claimed result. ■

Acknowledgment

We show our gratitude to the referees for their constructive comments, which help greatly to improve the presentation of this paper.

References

Baíllo, A. and Grané, A. (2009). Functional Local Linear Regression with Functional Predictor and Scalar Response. *Journal of Multivariate Analysis, **100**, 102-111.*

Barrientos-Marin, J., Ferraty, F. *and* Vieu, P. *(2010). Locally Modelled Regression and Functional Data.* J. of Nonparametric Statistics*, **22**, 617-632.*

Boj, E., Delicado, P. *and* Fortiana, J. *(2010). Distance-based local linear regression for functional predictors.* Computational Statistics and Data Analysis, **54**, 429-437.

Bosq, D. (2000). *Linear Processes in Function Spaces : Theory and applications. Lecture Notes in Statistics, **149**, Springer.*

Demongeot, J., Laksaci, A., Madani, F. *and* Rachdi, M. *(2011). A fast functional locally modeled conditional density and mode for functional time-series. Recent Advances in Functional Data Analysis and Related Topics, Contributions to Statistics, Physica-Verlag/Springer, 2011, Pages 85-90, DOI :* $10.1007/978-3-7908-2736-113$.

Demongeot, J., Laksaci, A., Madani, F. *and* Rachdi, M. *(2013). Functional data : local linear estimation of the conditional density and its application.*Statistics*. **47**, 26-44.*

Demongeot, J., Laksaci, A., Rachdi, M., *and* Rahmani, S. *(2014). On the local Modalization of the conditional distribution for functional data.* Sankhya A*, **76** (2), 328-355.*

Fan, J. *(1992). Design-adaptive nonparametric regression.* J. Amer. Statist. Assoc*. **87**, 998-1004.*

Fan, J. *and* Gijbels, I. *(1996).* Local Polynomial Modelling and its Applications. *Chapman & Hall, London.*

FAN, J. *and* YAO, Q. *(2003).* Nonlinear Time Series : Nonparametric and Parametric Methods. Springer-Verlag, New York.

FERRATY, F. and VIEU, P. (2006). *Nonparametric functional data analysis. Theory and Practice.* Springer Series in Statistics. New York.

FERRATY, F., LAKSACI, A., TADJ, A. and VIEU, P.(2010). Rate of uniform consistency for nonparametric estimates with functional variables, *Journal of Statistical Planning and Inference.* **140**, 335-352.

HONDA, T. (2000). Nonparametric estimation of a conditional quantile for α-mixing processes. *Annals of the Institute of Statistical Mathematics*, **52** (3), 459-470.

KOENKER, R. (2000). Galton, Edgeworth, Frisch, and prospects for quantile regression in econometrics. *Journal of Econometrics, **95**, 347-374.*

KOENKER, R. *(2005). A Frisch-Newton Algorithm for Sparse Quantile Regression.* Acta Mathematicae Applicatae Sinica, English Series, **21**, 225-236.

KOUL, H. L. and MUKHERJEE. K. (1994). Regression quantiles and related processes under long range dependent errors. *Journal of Multivariate Analysis*, **51**(2), 318-337.

PORTNOY, S. (1991). Correction: Asymptotic Behavior of M Estimators of p Regression Parameters when p^2/n is Large: II. Normal Approximation. *Ann. Statist, **19**, 2282.*

RACHDI, M., LAKSACI, A.,DEMONGEOT, J., ABDALI, A., MADANI, F. *(2014). Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data,* Computational Statistics and Data Analysis,***73**, 53-68.*

RAMSAY, J. O. *and* SILVERMAN, B.W. *(2002).* Applied functional data analysis: Methods and Case Studies. Springer Series in Statistics. Springer-Verlag, New York.

RAMSAY, J. O. and SILVERMAN, B.W. (2005). *Functional Data Analysis. 2nd ed.Springer, New-York.*

STONE, C. J. *(1977). Consistent Nonparametric Regression.* The Annals of Statistics, **5**(4), 595-620.

STUTE, W. (1986). Conditional Empirical Processes. *Ann. Statist*, **14**, 638-647.

USPENSKY, J.V. (1937). *Introduction to Mathematical Probability.* McGraw-Hill Book Company.

XIONG, X., ZHOU, P. and AILIAN, CH. (2018). Asymptotic normality of the local linear estimation of the conditional density for functional time series data, Communications in Statistics - Theory and Methods, **47**, 3418-3440 .

YU, K., LU, Z. and STANDER,J. (2003). Quantile regression: applications and current research areas. *Journal of the Royal Statistical Society Series D*, **52** (3), 331-350.

Chapter 5

Kernel estimation of mode regression for functional data.

For many years ago, a lot of papers are based on the Nadaraya-Watson kernel method for estimating the nonparametric regression function, then Ferraty and Vieu (2000) generalized the kernel regression estimator of Nadaraya-Watson familiar function to be more robust and flexible for functional data, this last model was adopted in many studies to find more asymptotic results.

When the merit of the derivatives estimator of regression is to provide us about the behaviour of both regression shape and mode regression, in view of that Mack and Muller (1989) used the Nadaraya-Watson kernel type to establish some consistency results such as : the asymptotic normality, the asymptotic mean squared error (AMSE) and the choice of bandwidth was discussed. M convergence of regression function estimator and its derivatives was studied by Boularan *et al.* (1995) and we point out that the reader can find important results in ρ-mixing case and α-mixing case by taking the monograph of Györfi *et al.* (1989). In mode regression framework, Ziegler (2002) established the probability convergence and the normality asymptotic of mode regression, this study is based on Nadaraya-Watson kernel estimator for scalar explanatory variable X and for the higher derivatives of regression estimator. For recent studies, we can see : Kemp and Santos Silva (2012) and Chen *et al.* (2017), with local linear method and for functional data, Hebchi (2020) introduced the model of mode regression and established the uniform almost complete convergence rate of local linear mode regression.

5.1 Model and assumptions

At this stage, we observe n pairs (X_i, Y_i) for $i = 1, ..., n$ identically distributed as (X, Y), this last is valued in $\mathcal{F} \times \mathbb{R}$, where $\mathcal{F}$ is a semi-metric space equipped with a semi-metric d. the functional kernel regression estimator $\hat{m}(.)$ is defined by :

$$\hat{m}(x) = \frac{\sum_{i=1}^{n} K(h^{-1}|\delta(x, X_i)|)Y_i}{\sum_{i=1}^{n} K(h^{-1}|\delta(x, X_i)|)} \tag{5.1}$$

where $\delta(.,.)$ is function defined from $\mathcal{F} \times \mathcal{F}$ to $\mathbb{R}$, such that: $d(.,.) = |\delta(.,.)|$. K is Kernel and $h = h_{K,n}$ is chosen as a sequence of positive real numbers and converges to 0 when $n \to \infty$.
with the convention $0/0 = 0$.
this chapter deals with mode regression estimator $\hat{\theta}$ of θ , where:

$$\hat{\theta}(x) = \sup_{x \in \mathcal{F}} \hat{m}(x) \tag{5.2}$$

whereas

$$\theta(x) = \sup_{x \in \mathcal{F}} m(x) \tag{5.3}$$

and on some subset $S_{\mathcal{F}}$ of $\mathcal{F}$

$$S_{\mathcal{F}} \subset \cup_{k=1}^{d_n} B(x_k, r_n)$$

where $x_k \in \mathcal{F}$ and r_n (resp d_n) is a sequence of positive real numbers.
To identify the asymptotic behaviour of our estimator, we need to take the following conditions:
the topological structure on the functional space $\mathcal{F}$ requires to use the following small-ball probability :

$$\forall \epsilon > 0, \quad \mathbb{P}(X \in B(x, \epsilon)) = \phi_x(\epsilon) > 0 \tag{5.4}$$

to establish pointwise convergence we assume that m has the continuity-type which defined by :

$$m \in \left\{ f : \mathcal{F} \times \mathbb{R}, \lim_{d(x,x') \to 0} f(x') = f(x) \right\} \tag{5.5}$$

The kernel K is a bounded and Lipschitz kernel on its support $[0, 1]$ and if its derivative K' exists on $[0, 1]$, we have for two real constants $-\infty < C_2 < C_1 < 0$:

$$C_2 \leq K' \leq C_1 \tag{5.6}$$

And $\phi_x(.)$ satisfies :

$$\exists C_3 > 0, \exists \eta_0 > 0, \forall 0 < \eta < \eta_0, \int_0^\eta \phi_x(u)du > C_3 \eta \phi_x(\eta) \tag{5.7}$$

the sequence of positive real numbers $(h_n)_{n \in \mathbb{N}}$ is such that

$$\lim_{n \to +\infty} h = 0, \quad \lim_{n \to +\infty} \frac{\log n}{n \phi_x(h)} = 0 \tag{5.8}$$

the scalar response variable Y satisfies

$$\forall k \geq 2, \ \mathbb{E}\left(|Y|^k | X = x\right) < \sigma_k(x) < \infty \text{ with } \sigma_k(.) \text{ continuous at } x \tag{5.9}$$

On the other side, the rate of convergence needs to add the following additional hypotheses which allow to precise the behaviour of the bias
The function ϕ satisfies

$$\forall x \in S_{\mathcal{F}}, \ 0 < C\phi(h) \leq \phi_x(h) \leq C'\phi(h) < \infty \tag{5.10}$$

The "Lipschitz-type" model is defined by : There exists $b > 0$ such that

$$\forall x_1, x_2 \in S_{\mathcal{F}}, |m(x_1) - m(x_2)| \leq Cd^b(x_1, x_2) \tag{5.11}$$

$m^{(j)}$ is j-times continuously differentiable around $\theta(x)$ with

$$m^{(l)}(\theta) = 0.(\text{for all } l = 1, ..., j-1) \text{ and } m^{(j)}(\theta) \neq 0 \tag{5.12}$$

for n large enough,

$$\frac{(\log n)^2}{n\phi(h)} < \log d_n < \frac{n\phi(h)}{\log n} \tag{5.13}$$

and

$$\sum_{n=1}^{\infty} \exp\{(1-\beta)\log d_n\} < \infty \text{ for some } \beta > 1. \tag{5.14}$$

5.2 Asymptotic results

Before given the asymptotic result,we introduce the following notations

$$\hat{m}_0(x) = \frac{1}{n\mathbb{E}[K(h^{-1}d(x,X_1))]}\sum_{i=1}^{n} K(h^{-1}d(x,X_1))$$

and

$$\hat{m}_1(x) = \frac{1}{n\mathbb{E}[K(h^{-1}d(x,X_1))]}\sum_{i=1}^{n} K(h^{-1}d(x,X_1))Y_i$$

and also the two importants theorems

Theorem 5.1 *under conditions (5.4)-(5.9)*

$$\lim_{n\to\infty}\sup_{x\in S_{\mathcal{F}}}|m(x)-\hat{m}(x)| = o(1), a.co.$$

Theorem 5.2 *(Ferraty et al. (2010)) under conditions (5.4), (5.6)-(5.14)*

$$\sup_{x\in S_{\mathcal{F}}}|\hat{m}(x)-m(x)| = O(h^b) + O\left(\sqrt{\frac{\ln d_n}{n\phi(h)}}\right) \quad a.co.$$

Theorem 5.3 *under equations (5.4) to (5.9) , we obtain that:*

$$\lim_{n\to+\infty}\hat{\theta}(x) = \theta(x), \quad a.co.$$

Theorem 5.4 *under conditions (5.4), (5.6)-(5.14)*

$$(\hat{\theta}(x)-\theta(x))^j = O(h^b) + O\left(\sqrt{\frac{\ln d_n}{n\phi(h)}}\right) \quad a.co.$$

5.3 A simulation study

In the sequel, we present a simulation study to see the performance of our estimator for functional data. so, for n, we generate our data by :

$$X_i(t_j) = \sin[(1 - W_i)t_j] + \sum_{k=1}^{j} \vartheta_{ik}, \text{for all } t \in (0, \pi) \;\; i = 1, 2, ...n, \text{ and } j = 1, ..., 100.$$

where : W_i is i.i.d. and follows the Normal distribution $\mathcal{N}(0, 1)$ and ϑ_{ik}'s are i.i.d. realizations of $\mathcal{N}(0, 0.1)$. In the following figure, for $n = 250$, we discretized our curves on the same grid which is composed of 100 equidistant values in $(0, \pi)$.

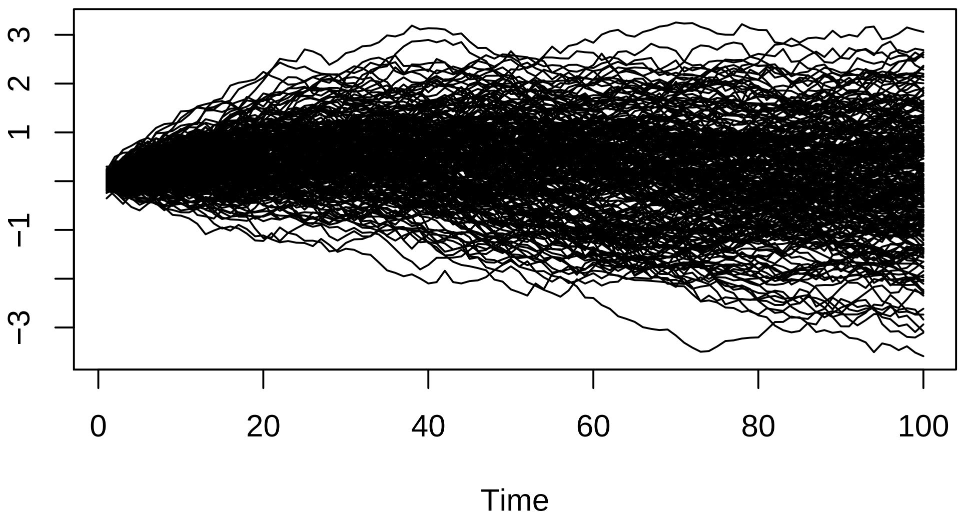

Figure 5.1: The curves Xi

The response sample is given by :

$$Y_i = m(X_i) + \epsilon_i$$

with : ϵ_i follows the Normal distribution $\mathcal{N}(0, 1)$. For the regression function, we take :

$$m(x) = \int_0^1 \frac{dt}{1 + |x(t)|}$$

and the mode regression is

$$\theta(x) = m(x) = \int_0^1 \frac{dt}{1+|x(t)|}$$

As our functional predictors are rough and for reducing dimensional space of data, we use the classical Principal Components Analysis (PCA) as a semi-metrics tool for computing proximities between curves (see Benhenni et al. (2007)). Then, we adopted the quadratic kernel (i.e. $K(u) = (3/4)(1-u^2)\mathbb{1}_{[0,1]}(u)$) as an asymmetrical kernel function. As is well known, the bandwidth selection plays a crucial role for the performance of a kernel estimate, hereinafter, we use the k-NN method with the same number of neighbours at any curve ($h(x) := h_{k_{\text{opt}}}(x)$), where the optimal number of neighbours obtained by :

$$k_{\text{opt}} = \arg\min_k \sum_{i=1}^{n} \left(Y_i - \theta_{(-i)}^{kNN}(X_i)\right)^2$$

where

$$\theta_{(-i)}^{kNN}(x) = \sup_{x\in S_{\mathcal{F}}} \hat{m}_{(-i)}^{kNN}(x)$$

with

$$\hat{m}_{(-i)}^{kNN}(x) = \frac{\displaystyle\sum_{j=1, j\neq i}^{n} K(h_k^{-1}(x)|\delta(x,X_j)|)Y_j}{\displaystyle\sum_{j=1, j\neq i}^{n} K(h_k^{-1}(x)|\delta(x,X_j)|)}$$

in order to examine the performance of mode regression estimator, we split randomly our initial sample into two subsample :learning sample ($\mathcal{L}$) and testing sample ($\mathcal{T}$), where : $\mathcal{L}\cap\mathcal{T} = \emptyset$ and we take : $|\mathcal{L}| = 200$, $|\mathcal{T}| = 50$. Then, we end our study by computing the mean square errors (MSE) of prediction in order to highlight the performance of our method :

$$MSE(\mathcal{L},\mathcal{T}) = \frac{1}{|\mathcal{T}|}\sum_{i\in\mathcal{T}}(m(X_i) - \hat{\theta}(X_i))^2 = 0.107.$$

Finally, we conclude that our estimator gives better estimation results.

5.4 Appendix

Proof of theorem 5.1

Remark that, the theoreme's proof can be deduced directly from the following decomposition

$$\begin{aligned}\hat{m}(x) - m(x) &= \tfrac{1}{\hat{m}_0(x)}\left\{(\hat{m}_1(x) - \mathbb{E}[\hat{m}_1(x)]) + (\mathbb{E}[\hat{m}_1(x)] - m(x))\right\} \\ &\quad + \tfrac{m(x)}{\hat{m}_0(x)}\left(1 - \hat{m}_0(x)\right)\end{aligned} \tag{5.15}$$

theorem 5.1 is a result of the following intermediate lemmas.

Lemma 5.1 *under the conditions of theorem 5.1*

$$\lim_{n\to\infty} \sup_{x\in S_{\mathcal{F}}} |\hat{m}_0(x) - 1| = O\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right), \ a.co.$$

Corollary 5.1

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{x\in S_{\mathcal{F}}} \hat{m}_0(x) < \frac{1}{2}\right) < \infty$$

Lemma 5.2 *under the conditions of theorem 5.1*

$$\lim_{n\to\infty} \sup_{x\in S_{\mathcal{F}}} |\hat{m}_1(x) - \mathbb{E}[\hat{m}_1(x)]| = O\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right). \ a.co.$$

Lemma 5.3 *under the conditions of theorem 5.1*

$$\lim_{n\to\infty} \sup_{x\in S_{\mathcal{F}}} |\mathbb{E}[\hat{m}_1(x)] - m(x)| = o(1), \ a.co.$$

As in Ferraty*et al.* (2010) (see lemma 8 and lemma 11), we can prove lemma(5.1) and lemma(5.2) such that we use the condition (5.5) instead of condition $H2$ in Ferraty*et al.* (2010) and $\Psi_{S_{\mathcal{F}}}\left(\frac{\log n}{n}\right) = O(\log n)$

Proof of corollary 5.1

$\inf_{x\in S_{\mathcal{F}}} |\hat{m}_0(x)| \leq \frac{1}{2} \Longrightarrow \exists x \in S_{\mathcal{F}}, 1 - \hat{m}_0(x) \geq \frac{1}{2} \Longrightarrow \sup_{x\in S_{\mathcal{F}}} |1 - \hat{m}_0(x)| \geq \frac{1}{2}$

from lemma 5.1 we can deduce that

$$\mathbb{P}\left(\inf_{x\in S_{\mathcal{F}}} \hat{m}_0(x) < \frac{1}{2}\right) \leq \mathbb{P}\left(\sup_{x\in S_{\mathcal{F}}} |1 - \hat{m}_0(x)| > \frac{1}{2}\right) < \infty$$

■

Proof of lemma 5.3

$$\begin{aligned}|\mathbb{E}[\hat{m}_1(x)] - m(x)| &= |\frac{1}{n\mathbb{E}[K(h^{-1}d(x,X_1))]}\mathbb{E}[\sum_{i=1}^{n} K(h^{-1}d(x,X_1))Y_i] - m(x)| \\ &= |\frac{1}{\mathbb{E}[K(h^{-1}d(x,X_1))]}\mathbb{E}[K(h^{-1}d(x,X_1))Y_1] - m(x)| \\ &\leq \frac{1}{\mathbb{E}[K(h^{-1}d(x,X_1))]}\mathbb{E}[K(h^{-1}d(x,X_1))|m(X_1) - m(x)|] \\ &\leq \sup_{x' \in S_{\mathcal{F}}} |m(x) - m(x')|\end{aligned}$$

by using (5.5) we get the claimed result. ■

Proof of theorem 5.3 if θ is the unique solution of (5.3), we have that $\forall \epsilon > 0, \exists \mu > 0$, such that

$$|\theta - x| \geq \epsilon \Longrightarrow |r(\theta) - r(x)| \geq \mu \text{ (for all } x \in S_{\mathcal{F}})$$

on the other hand, the definitions of $\theta(x)$ and $\hat{\theta}(x)$ lead us to

$$\begin{aligned}|m(\theta(x)) - m(\hat{\theta}(x))| &= |m(\theta(x)) - \hat{m}(\theta(x)) + \hat{m}(\theta(x)) - m(\hat{\theta}(x))| \\ &\leq |m(\theta(x)) - \hat{m}(\theta(x))| + |\hat{m}(\theta(x)) - m(\hat{\theta}(x))| \\ &\leq 2\sup_{x \in S_{\mathcal{F}}} |m(x) - \hat{m}(x)|\end{aligned}$$

by using theorem (5.1) we get

$$\forall \epsilon > 0, \sum_{n=1}^{\infty} \mathbb{P}(|\hat{\theta}(x) - \theta(x)| > \epsilon) < \infty.$$

■

Proof of theorem 5.4 due to Taylor expanssion of oder j of the function m at point $\theta(x)$ and from equation (5.12), we have

$$m(\hat{\theta}(x)) = m(\theta(x)) + \frac{1}{j!}m^{(j)}(\theta')(\theta(x) - \hat{\theta}(x))^j$$

and as

$$|m(\theta(x)) - m(\hat{\theta}(x))| \leq 2 \sup_{x \in S_{\mathcal{F}}} |m(x) - \hat{m}(x)|$$

by combining the two latter results, we can write

$$m^{(j)}(\theta')[\theta(x) - \hat{\theta}(x)]^j = O\left(\sup_{x \in S_{\mathcal{F}}} |\hat{m}(x) - m(x)|\right) \tag{5.16}$$

theorem (5.3) insures that

$$\lim_{n \to \infty} m^{(j)}(\theta') = m^{(j)}(\theta(x)) \neq 0. \tag{5.17}$$

by using (5.16), (5.17) and Proposition A.6-ii in Ferraty and Vieu (2006), we obtain

$$|\theta(x) - \hat{\theta}(x)|^j \leq C \sup_{x \in S_{\mathcal{F}}} |\hat{m}(x) - m(x)|, \ \ a.co.$$

this last inequality together with theorem 5.2, we get the claimed result. ■

References

Benhenni, K., Ferraty, F., Rachdi, M. and Vieu, P. (2007). Local smoothing regression with functional data. *Computational Statistics. **22**(3), 353- 369.*

Boularan, J., Ferré, L. and Vieu, P. *(1995). Location of particular points in nonparametric regression analysis.* Austral. J. Statist. ***37**, 161-168.*

Chen, Y., Ma, X. *and* Zhou, J. *(2017). Variable selection for mode regression.* Journal of Applied Statistics, **45**, 1-8.

Ferraty, F. and Vieu, P. (2000). Dimension fracale et estimation de la régression dans des espaces vectoriels semi-normés, *Compte Rendus de l'Académie des Sciences Paris, **330**, 403-406.*

Ferraty, F. and Vieu, P. (2006). Nonparametric functional data analysis. Theory and Practice. *Springer Series in Statistics. New York.*

Ferraty, F., Laksaci, A., Tadj, A., *and* Vieu, P.*(2010). Rate of uniform consistency for nonparametric estimates with functional variables,* Journal of Statistical Planning and Inference. ***140**, 335-352.*

Györfi, L., Härdle, W., Sarda, P. *and* Vieu,P. *(1989).* Nonparametric Curve Estimation from Time Series. Springer-Verlag New York.

Hebchi, C. (2020). Uniform almost complete convergence of local linear mode regression, *IJSE, **21**(1), 54-62.*

Kemp, G. *and* Santos Silva, J. *(2012). Regression towards the mode.*Journal of Econometrics, **170**, 92-101.

Mack ,Y. P. and Müller H. G. (1989). Derivative Estimation in Nonparametric Regression with Random Predictor Variable.*Sankhyā, **51**(1), 59-72.*

Ziegler, K. *(2002). On nonparametric kernel estimation of the mode of the re-*

gression function in the random design model. Journal of Nonparametric Statistics, **14**(6), 749-774.

Chapter 6

Uniform almost complete convergence of local linear mode regression

This chapter is published in International Journal of Statistics and Economics, (2020), **21**(1), 54-62.

Uniform almost complete convergence of local linear mode regression

Chaima HEBCHI
Lab. of Statistics and Stochastic Process (LSPS)
Djillali liabès University
BP 89, Sidi bel Abbes 22000 Algeria
chaimahabchi@yahoo.fr

Abstract

The aim of this paper is to join the advantages of mode with regression function by using local linear method in order to establish the Uniform almost complete convergence in iid setting and for functional data.
Key Words: functional data, locally modelled regression, mode regression.
2000 Mathematics Subject Classification: 62G05, 62G08, 62G20.

6.1 Introduction

The mode is a major contributor to study the link between a response variable and preductors and also to provide an important summary of data. As the technolology's advence easies and facilitates the collecting and storing data in high-dimension, the mode has long been a question of great interest in a wide range of fields such as : biology, astronomy and econometrics and others.

This paper is concerned with mode regression, for old studies and in regression framework we can see (Rousseeuw, 1984) and (Powell, 1986). In 1993, (Lee, 1993) generalized the mode regression with rectangular kernel (RME) to quadratic kernel (QME) where RME and QME are suitable when the dependent variable is truncated

Recently, (Kemp and Santos Silva, 2012) stated the probability convergence and asymptotic normality of mode regression based on the standard normal density kernel with (Parzen, 1962) regression version of mode estimator (semi-parametric mode regression estimator). In nonparametric kernel estimation, (Yao and Li,

2013) proposed the MEM (Modal Expectation Maximization) algorithm to estimate the modal regression parameter, furthermore, under some technical conditions they estitablished convergence rate and the asymptotic normality of the modal regression coefficient. In high-dimensional data with nonparametric kernel estimation, (Chen, Ma and Zhou, 2017) used SCAD (Smoothly Clipped Absolute Deviation) penalty for variables selection and by extending MEM, they proposed PMEM (Penalized Model Expectation Minimization) algorithm to estimate the parameters of mode regression and they exploited all these to prove the consistency and the sparsity of the resultant estimator.

More than decade, researchers have shown an increased interest in functional data analysis, we refer to the the monographs of (Ramsay and Silverman, 2002), (Ramsay and Silverman, 2005), (Bosq, 2000) and (Ferraty and Vieu, 2006)for more theories and applications and as method we use the local linear which has various adventages see (Fan and Gijbels, 1992), (Baíllo and Grané, 2009) and (Barrientos-Marin, Ferraty and Vieu, 2010).

In the iid setting, our work focus on the local linear estimation of the mode regression for functional data. In section 2, This paper begins by clarifying our model and proposing our estimators, under some assumptions the main results are stated in section 3,Proofs can be found in section 4.

6.2 Model Framework And Conditions

6.2.1 Model

we consider n pairs $(X_i, Y_i)_{i=1,\dots,n}$ identically and independently distributed as (X, Y), this last is valued in $\mathcal{F} \times \mathbb{R}$, where $\mathcal{F}$ is a semi-metric space equipped with a semi-metric d
The mode of regression function θ on $\mathcal{F}$ is

$$\theta(x) = \sup_{x \in \mathcal{F}} m(x)$$

whereas, the mode regression estimator $\hat{\theta}$ is defined by

$$\hat{\theta}(x) = \sup_{x \in \mathcal{F}} \hat{m}(x)$$

where : m and $\hat{m}$ are regression and regression estimator respectively. To get the estimator of local linear regression we minimise the following quantity :

$$\min_{(a,b)\in\mathbb{R}^2} \sum_{i=1}^{n} |Y_i - a - b\beta(X_i, x))|^2 K(h_K^{-1}\delta(X_i, x)) \tag{6.1}$$

where $\beta(.,.)$ and $\delta(.,.)$ are two functions defined from $\mathcal{F} \times \mathcal{F}$ to $\mathbb{R}$, such that:

$$\forall \xi \in \mathcal{F}, \beta(\xi, \xi) = 0, \text{and} \quad d(.,.) = |\delta(.,.)|.$$

K is Kernel and $h_K = h_{K,n}$ is chosen as a sequence of positive real numbers. by a simple algebra we define explicitly the regression function estimator $\hat{m}$ and its first derivative $\hat{m}^{(1)}$:

$$\begin{pmatrix} \hat{m} \\ \hat{m}^{(1)} \end{pmatrix} = \frac{1}{\sum_{i,j=1}^{n} \beta_i(\beta_i - \beta_j) K_i K_j} \begin{pmatrix} \sum_{i,j=1}^{n} \beta_i(\beta_i - \beta_j) K_i K_j Y_j \\ \sum_{i,j=1}^{n} (\beta_j - \beta_i) K_i K_j Y_j \end{pmatrix}$$

6.2.2 Assumptions

our study aims to state the uniform almost-complete convergence of $\hat{\theta}$ on some subset $S_{\mathcal{F}}$ of $\mathcal{F}$ such that

$$S_{\mathcal{F}} \subset \cup_{k=1}^{d_n} B(x_k, r_n)$$

where $x_k \in \mathcal{F}$ and r_n (resp d_n) is a sequence of positive real numbers, and we suppose that our estimator satisfies the following conditions :

$(A1)$ For any $r > 0$, $\phi_x := \mathbb{P}(x \in B(x, r))$ and there is a function τ_x such that:

$$\forall t \in [0,1], \lim_{h \to 0} \frac{\phi(th, h)}{\phi(h)} = \tau_x(t)$$

$$\phi(r_1, r_2) = \mathbb{P}(r_1 < |\delta(x, x')| < r_2)$$

$(A2)$ m is twice continuously differentiable on $S_{\mathcal{F}}$ with bounded second order derivatives.

$(A3)$ $m \in \{f : \mathcal{F} \to \mathbb{R}, \exists C \in \mathbb{R}_+, x' \in \mathcal{F}, |f(x) - f(x')| < C|\delta(x, x')|^b\}$

$(A4)$ To control the shape of local functional object β we have :

$$\exists 0 < M_1 < M_2, \forall x' \in \mathcal{F}, M_1|\delta(x, x')| \leq |\beta(x, x')| \leq M_2|\delta(x, x')|$$

$(A5)$ The kernel function Ksatisfies : K is a positive, differentiable function with support $[0, 1]$

$(A6)$ The local expectation of β satisfies :

$$h \int_{B(x,h)} \beta(u, x)dP(u) = o(\int_{B(x,h)} \beta^2(u, x)dP(u))$$

$(A7)$ $\forall k \geq 1, \mathbb{E}[|Y|^k|X = x] < C$

$(A8)$ $\forall x' \in \mathcal{F} : m(x') = m(x) + \beta(x, x')m^{(1)}(x)$

$(A9)$ The behaviour of the bandwidth h is :

$$\exists n_0, \forall n > n_0, \frac{1}{\phi(h)} \int_0^1 \phi(Zh, h) \frac{d}{dZ}(Z^j K(Z))dZ > c > 0, \ j = 1, 2$$

$(A10)$ for $r_n = O\left(\frac{\ln n}{n}\right)$ the sequence d_n statisfies :

$$\frac{(\ln n)^2}{n} < \ln d_n < \frac{n\phi_x(h)}{\ln n}$$

and

$$\exists \beta > 1 : \sum_{i=1}^{\infty} d_n^{1-\beta} < \infty$$

we remark that all of these conditions are commonly used in many studies of the local linear method for functional data.

6.3 Asymptotic results

Theorem 6.1 *under assumptions* $(A1) - (A10)$ *we get*

$$\sup_{x \in S_{\mathcal{F}}} |\hat{m}^{(1)}(x) - m^{(1)}(x)| = O(h^b) + O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\phi_x(h)}}\right)$$

Proof of theorem6.1
with a Taylor development of $\hat{m}^{(1)}(\hat{\theta}(x))$ around $\theta(x)$, we get:

$$\hat{\theta} - \theta = \frac{1}{\hat{m}^2(\theta^*)}(\hat{m}^{(1)}(\theta) - m^{(1)}(\theta))$$

the unimodality of m and under the condition $(A2)$ we have

$$m^{(1)}(\theta(x)) = \hat{m}^{(1)}(\theta(x)) = 0 \text{ and} m^{(2)}(\theta(x)) < 0$$

by using the result of Proposition 3.5 in (Pons, 2011), with the fact that : $\theta^*(x)$ is lying between $\hat{\theta}(x)$ and $\theta(x)$ and by : $\hat{\theta}(x) - \theta(x) \to 0_{a.co.}$ (see Theorem 6.6 in book of (Ferraty and Vieu, 2006) and (Pons, 2011) p.68) , it follows that

$$\hat{m}^2(\theta^*(x)) - m^2(\theta(x)) \to 0_{a.co.}$$

Since, $|m^2(\theta(x))| > 0$, we can see :

$$\exists C > 0, \text{such that} |\frac{1}{\hat{m}^2(\theta^*(x))}| \leq C_{a.co.}$$

by remarking that
$\hat{m}^{(1)}(x) - m^{(1)}(x) =$

$$\frac{1}{\hat{m}_0^{(1)}(x)}\left\{ \left(\hat{m}_1^{(1)}(x) - \mathbb{E}(\hat{m}_1^{(1)}(x))\right) - \left(m^{(1)}(x)\mathbb{E}(\hat{m}_0^{(1)}(x)) - \mathbb{E}(\hat{m}_1^{(1)}(x))\right) \right\}$$

$$-\frac{m^{(1)}(x)}{\hat{m}_0^{(1)}(x)}\left\{\hat{m}_0^{(1)}(x) - \mathbb{E}(\hat{m}_0^{(1)}(x))\right\}$$

where:

$$\hat{m}_1^{(1)}(x) = \frac{1}{n(n-1)\mathbb{E}[\beta_1 K_1 K_2]}\sum_{i,j=1}^{n}(\beta_j - \beta_i)K_iK_jY_j$$

$$\hat{m}_0^{(1)}(x) = \frac{1}{n(n-1)\mathbb{E}[\beta_1 K_1 K_2]}\sum_{i,j=1}^{n}\beta_i(\beta_j - \beta_i)K_iK_j$$

the theoreme's proof can be deduced directly from the following lemmas

Lemma 6.1 *under $(A1)$, $(A2)$ and $(A7)$ we have:*

$$\sup_{x \in S_{\mathcal{F}}} |m^{(1)}(x)\mathbb{E}(\hat{m}_0^{(1)}(x)) - \mathbb{E}(\hat{m}_1^{(1)}(x))| = O(h^b)$$

Lemma 6.2 *under the hypotheses* $(A1)$, $(A3)-(A9)$ *we obtain that:*

$$\sup_{x\in S_{\mathcal{F}}} |\hat{m}_1^{(1)}(x) - \mathbb{E}(\hat{m}_1^{(1)}(x))| = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\phi_x(h)}}\right)$$

Lemma 6.3 *under* $(A1)$, $(A3)-(A9)$ *we obtain*

$$\sup_{x\in S_{\mathcal{F}}} |\hat{m}_0^{(1)}(x) - \mathbb{E}(\hat{m}_0^{(1)}(x))| = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\phi_x(h)}}\right)$$

Corollary 6.1 *under conditoins of 6.3,* $\exists C_3 > 0$ *such that*

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{x\in S_{\mathcal{F}}} \hat{m}_0^{(1)}(x) < C_3\right) < \infty$$

6.4 Appendix

Proof of Lemma 6.1 first of all, we have :

$$\begin{aligned}|m^{(1)}(x)\mathbb{E}(\hat{m}_0^{(1)}(x)) - \mathbb{E}(\hat{m}_1^{(1)}(x))| =& \frac{1}{|\mathbb{E}(\beta_1 K_1 K_2)|}\Big|\mathbb{E}(\beta_1(\beta_1-\beta_2)K_1K_2m^{(1)}(x)\\ &- (\beta_2-\beta_1)K_1K_2m(X_2))\Big|\end{aligned}$$

by using $(A7)$ we can see:

$$m(x_1) - m(x_2) = (\beta_1 - \beta_2)m^{(1)}(x)$$

and we obtain
$|m^{(1)}(x)\mathbb{E}(\hat{m}_0^{(1)}(x)) - \mathbb{E}(\hat{m}_1^{(1)}(x))|$

$$= \frac{1}{|\mathbb{E}(\beta_1 K_1 K_2)|}\left|\mathbb{E}\left(\beta_1(m(x_1)-m(x_2))K_1K_2-(\beta_2-\beta_1)m(x_2)K_1K_2\right)\right|$$

$$= \frac{1}{|\mathbb{E}(\beta_1 K_1 K_2)|}\left|\mathbb{E}\left(\beta_1 m(x_1)K_1K_2-\beta_2 m(x_2)K_1K_2\right)\right|$$

$$= \frac{1}{|\mathbb{E}(\beta_1 K_1 K_2)|}\left|\mathbb{E}\left(\beta_1 m(x_1)K_1K_2-\beta_1 m(x)K_1K_2+\beta_1 m(x)K_1K_2\right.\right.$$

$$\left.\left.-\beta_2 m(x_2)K_1K_2+\beta_2 m(x)K_1K_2-\beta_2 m(x)K_1K_2\right)\right|$$

$$\leq \frac{1}{|\mathbb{E}(\beta_1 K_1 K_2)|}\left|\mathbb{E}[\beta_1(m(x_1)-m(x))K_1K_2]\right|$$

$$+\frac{1}{|\mathbb{E}(\beta_1 K_1 K_2)|}\left|\mathbb{E}[\beta_2(m(x)-m(x_2))K_1K_2]\right|$$

we are in position to use the assumption $(A3)$: such that

$$\mathbb{1}_{B(x,h)}|m(x)-m(x_1)| \leq Ch^b$$

$$\mathbb{1}_{B(x,h)}|m(x)-m(x_2)| \leq Ch^b$$

and we get

$$\sup_{x\in S_{\mathcal{F}}} |m^{(1)}(x)\mathbb{E}(\hat{m}_0^{(1)}(x))-\mathbb{E}(\hat{m}_1^{(1)}(x))| \leq Ch^b$$

■

Proof of Lemma 6.2 one starts the proof by considering the following decomposition

$$\hat{m}_1^{(1)}(x) = \underbrace{\frac{n^2 h\phi^2}{n(n-1)\mathbb{E}(\beta_1K_1K_2)}}_{T}\left[\underbrace{\left(\frac{1}{n}\sum_{j=1}^{n}\frac{\beta_jK_jY_j}{h\phi_x(h)}\right)}_{S_1(x)}\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{K_i}{\phi_x(h)}\right)}_{S_2(x)}\right.$$

$$\left.-\underbrace{\left(\frac{1}{n}\sum_{j=1}^{n}\frac{K_jY_j}{\phi_x(h)}\right)}_{S_3(x)}\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{K_i\beta_i}{h\phi_x(h)}\right)}_{S_4(x)}\right]$$

$\hat{m}_1^{(1)}(x)-\mathbb{E}[\hat{m}_1^{(1)}(x)] = T[S_1(x)S_2(x)-\mathbb{E}(S_1(x)S_2(x))-(S_3(x)S_4(x)-\mathbb{E}(S_3(x)S_4(x))]$ it remains to prove:

$$T = O(1) \tag{6.2}$$

$$\sup_{x\in S_{\mathcal{F}}} |S_i - \mathbb{E}[S_i]| = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\phi_k(x)}}\right) \text{ for } i = 1,2,3,4 \tag{6.3}$$

$$\sup_{x\in S_{\mathcal{F}}} |\mathbb{E}[S_1]\mathbb{E}[S_2] - \mathbb{E}[S_1S_2] - \mathbb{E}[S_3]\mathbb{E}[S_4] + \mathbb{E}[S_3S_4]| = o_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\phi_x(h)}}\right) \tag{6.4}$$

Proof of (6.2) : by using $(A4)$, it is easy to see that :

$$\mathbb{E}[K_1\beta_1] > M_1\mathbb{E}[K_1 d(x_1, x)]$$

Moreover, we can write

$$\begin{aligned}
\mathbb{E}[K_1\tfrac{d(x_1,x)}{h}] &= \int_0^1 tK_1(t)dP^{\frac{d(x_1,x)}{h}}(t) \\
&= \int_0^1 \left[\int_0^t (\tfrac{d}{du}(uK_1(u)))du\right] dP^{\frac{d(x_1,x)}{h}}(t) \\
&= \int_0^1 \left[(\int_0^1 \mathbb{1}_{[u,1]}(t)dP^{\frac{d(x_1,x)}{h}}(t))\tfrac{d}{du}(uK_1(u))\right] du \\
&= \int_0^1 \phi_x(uh,h)\tfrac{d}{du}(uK(u))du \\
&= \phi_x(h)[\tfrac{1}{\phi_x(h)}\int_0^1 \phi_x(uh,h)\tfrac{d}{du}(uK(u))du]
\end{aligned}$$

by using $(A9)$, we get

$$\mathbb{E}[K_1\beta_1] > M_1 h\phi_x(h) \tag{6.5}$$

which leads us to: $T = O(1)$

Proof of (6.3) : we note $j(x) = \arg\min_{j\in\{1,2,\dots,d_n\}} |\delta(x, x_k)|$ to consider the following decomposition

$$\begin{aligned}
\sup_{x\in S_{\mathcal{F}}} |S_i(x) - \mathbb{E}[S_i(x)]| \leq\ & \underbrace{\sup_{x\in S_{\mathcal{F}}} |S_i(x) - S_i(x_{j(x)})|}_{T_1} + \underbrace{\sup_{x\in S_{\mathcal{F}}} |S_i(x_{j(x)}) - \mathbb{E}[S_i(x_{j(x)})]|}_{T_2} \\
& + \underbrace{\sup_{x\in S_{\mathcal{F}}} |\mathbb{E}[S_i(x_{j(x)})] - \mathbb{E}[S_i(x)]|}_{T_3}
\end{aligned}$$

Treatment of T_1 and T_3 : by the same treatment of F_1^k in (Demongeot, Laksaci, Madani and Rachdi, 2013) we get :

$$T_1 = O\left(\sqrt{\frac{\ln d_n}{n\phi_k(x)}}\right), \quad T_3 = O\left(\sqrt{\frac{\ln d_n}{n\phi_k(x)}}\right)$$

Treatment of T_2 : we set:

$$\Gamma_i^{k,l} = \frac{1}{h^k\phi_x(h)}\left(K_i\beta_i^kY_i^l - \mathbb{E}[K_i\beta_i^kY_i^l]\right), \text{ for } i=1,2,3,4,\ k=0,1 \text{ and } l=0,1$$

By Newton's binomial expansion, we obtain:

$$\begin{aligned}\mathbb{E}\left(|K_i\beta_i^kY_i^l - \mathbb{E}[K_i\beta_i^kY_i^l]|^m\right) &= \mathbb{E}|\sum_{d=0}^{m} C_m^d (K_i\beta_i^kY_i^l)^d(\mathbb{E}[K_i\beta_i^kY_i^l])^{m-d}(-1)^{m-1}| \\ &\leq \sum_{d=0}^{m} C_m^d \mathbb{E}(|K_i\beta_i^kY_i^l|^d)|\mathbb{E}[K_i\beta_i^kY_i^l]|^{m-d} \\ &\leq \sum_{d=0}^{m} C_m^d \mathbb{E}|K_1^d\beta_1^{kd}\mathbb{E}[|Y_1^l|^d|X_1]||\mathbb{E}[K_1\beta_1^k\mathbb{E}[|Y_1^l|^d|X_1]]|^{m-d}\end{aligned}$$

where $C_m^d = m!/(d!(m-d)!)$
The condition $(A7)$ allows us to show that:
for $i=1,2,3,4,\ k=0,1$ and $l=0,1$

$$\mathbb{E}|\Gamma_i^{k,l}|^m = O(\phi_x^{-m+1}(h))$$

by using the classical Bernstein's inequality, we get:

$$P\left(|\sum_{i=1}^{n}\Gamma_i^{k,l}| > \eta\sqrt{\frac{\ln d_n}{n\phi_k(x)}}\right) \leq 2d_n^{-C\eta^2}$$

this last equation leads us to get:

$$\forall j \leq d_n,\quad P\left(|S_i(x_j) - \mathbb{E}[S_i(x_j)]| > \eta\sqrt{\frac{\ln d_n}{n\phi_k(x)}}\right) \leq 2\exp\{-C\eta^2\ln d_n\}$$

Since, $\frac{\ln d_n}{n\phi_k(x)} > \frac{1}{n\phi_k(x)}$ and by choosing η such that: $C\eta^2 = \beta$ we can see:

$$d_n \max_{j\in\{1,\ldots,d_n\}} P\left(|S_i(x_j) - \mathbb{E}[S_i(x_j)]| > \eta\sqrt{\frac{\ln d_n}{n\phi_x(h)}}\right) \leq C' d_n^{1-\beta}$$

and the second part of $(A10)$ allows us to get the claimed result.

Proof of (6.4) : to prove equation (6.4) we use the fact that the pairs $(X_i, Y_i)_{i=1,\dots,n}$ are identically distributed as (X, Y), by using the Lemma $A.1$ in (Barrientos-Marin et al., 2010) and $(A7)$ we have: $\mathbb{E}[K_i\beta_i Y_i] = O(h\phi_x(h))$
which implies that:

$$\sup_{x\in S_{\mathcal{F}}} |\mathbb{E}[S_1]\mathbb{E}[S_2] - \mathbb{E}[S_1S_2] - \mathbb{E}[S_3]\mathbb{E}[S_4] + \mathbb{E}[S_3S_4]| = o_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\phi_x(h)}}\right)$$

■

Proof of Lemma 6.3 the definition of $\hat{m}_0^{(1)}$ leads us to the following decomposition :

$$\hat{m}_0^{(1)} = \underbrace{\frac{n^2h^2\phi^2}{n(n-1)\mathbb{E}(\beta_1K_1K_2)}}_{T'} \left[\underbrace{\left(\frac{1}{n}\sum_{j=1}^{n}\frac{K_j}{\phi_x(h)}\right)}_{S_1'(x)}\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{K_i\beta_i^2}{h^2\phi_x(h)}\right)}_{S_2'(x)} - \underbrace{\left(\frac{1}{n}\sum_{j=1}^{n}\frac{K_j\beta_j}{h\phi_x(h)}\right)}_{S_3'(x)}\underbrace{\left(\frac{1}{n}\sum_{j=1}^{n}\frac{K_j\beta_j}{h\phi_x(h)}\right)}_{S_4'(x)}\right]$$

all it remains to prove are the following equations :

$$T' = O(1) \tag{6.6}$$

and $\forall x \in S_{\mathcal{F}}$ and for $i = 1, 2, 3, 4$

$$\sup_{x\in S_{\mathcal{F}}} |S_i' - \mathbb{E}[S_i']| = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\phi_x(h)}}\right) \text{ and } |\mathbb{E}[S_i']| = O(1) \tag{6.7}$$

and, also that

$$\sup_{x\in S_{\mathcal{F}}} |\mathbb{E}[S_1']\mathbb{E}[S_2'] - \mathbb{E}[S_1'S_2'] - \mathbb{E}[S_3']\mathbb{E}[S_4'] + \mathbb{E}[S_3'S_4']| = o_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\phi_x(h)}}\right) \tag{6.8}$$

due to (6.5) and $h < 1$ we get $T' = O(1)$
we remarque that equations (6.7) and (6.8) have the same arguments as in Lemma 3.4.1 of (Demongeot et al., 2013) which implies the same treatments ■

Proof of Corollary 6.1 we know that : $0 < h < 1$ and $\mathbb{E}[\beta_1 K_1] \leq Ch\phi_x(h)$ under assumption $(A4)$ we have : $\mathbb{E}[\beta_1^2 K_1] \geq \mathbb{E}[\delta_1^2 K_1]$
the same steps as in lemma A.1(ii) (see (Barrientos-Marin et al., 2010)) leads us to obtain

$$\frac{\mathbb{E}[\beta_1^2 K_1]}{h^2\phi_x(h)} > C > 0$$

at this stage we can show : there exists a real number $C^{'}$ such that

$$\mathbb{E}[\hat{m}_0^{(1)}(x)] \geq C^{'} \text{for all } x \in S_{\mathcal{F}}$$

hence, $\inf_{x \in S_{\mathcal{F}}} \hat{m}_0^{(1)}(x) \geq \frac{C^{'}}{2} \Longrightarrow \exists x \in S_{\mathcal{F}} \text{such that} : |\mathbb{E}[\hat{m}_0^{(1)}(x)] - \hat{m}_0^{(1)}(x)| \geq \frac{C^{'}}{2}$
which leads to :

$$\sup_{x \in S_{\mathcal{F}}} |\mathbb{E}[\hat{m}_0^{(1)}(x)] - \hat{m}_0^{(1)}(x)| \geq \frac{C^{'}}{2}$$

by using lemma 6.3 and for $C_3 = \frac{C^{'}}{2}$ we get

$$\sum_n \mathbb{P}\left(\inf_{x \in S_{\mathcal{F}}} (x) \leq C_3 \right) \leq \sum_n \mathbb{P}\left(\sup_{x \in S_{\mathcal{F}}} |\mathbb{E}[\hat{m}_0^{(1)}] - \hat{m}_0^{(1)}| \geq C_3 \right) < \infty$$

■

Acknowledgment

we would like to thank the editor and reviewer for their valuable suggestions and discussions.

References

BAÍLLO, A. and GRANÉ, A. (2009). Functional Local Linear Regression with Functional Predictor and Scalar Response. *Journal of Multivariate Analysis, **100**, 102-111.*

BARRIENTOS-MARIN, J., FERRATY, F. *and* VIEU, P *.(2010). Locally Modelled Regression and Functional Data.*J. of Nonparametric Statistics, **22**, 617-632.

BOSQ., D .(2000). *Linear Processes in Function Spaces: Theory and Applications.* Springer-Verlag.

CHEN, Y., MA, X. and ZHOU, J. (2017). Variable selection for mode regression. *Journal of Applied Statistics, **45**, 1-8 .*

DEMONGEOT, J., LAKSACI, A., MADANI, F. *and* RACHDI, M. *(2013). Functional data : local linear estimation of the conditional density and its application.J. of Statistics, **47**, 26-44.*

FAN, J. *and* GIJBELS, I. *(1992). Variable Bandwidth and Local Linear Regression Smoothers. Annals of Statistics, **20**, 2008-2036.*

FERRATY, F. *and* VIEU, P *. (2006). Nonparametric Functional Data Analysis : Theory and Practice.Springer-Verlag New York.*

KEMP, G.C.R. *and* SANTOS SILVA,J.M.C. *(2012). Regression towards the mode.Journal of Econometrics, **170**, 92-101.*

LEE, M.J. *(1989). Mode regression.J. Econometrics, **42**, .*

LEE, M.J.. *(1993). Quadratic mode regression.*J. Econometrics, **57**, 1-19 .

PARZEN, E. (1962). Estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, **33**, 1065-1076.

PONS, O. (2011). *F*unctional Estimation for Density, Regression Models and Processes. WORLD SCIENTIFIC, Singapore.

POWELL, J. (1986). Symmetrically trimmed least squares estimation for Tobit models. *Econometrica*, **54**, 1435-1460.

ROUSSEEUW, P. J. (1984). Least Median of Squares Regression.*Journal of the American Statistical Association*, **79**, 871-880 .

RAMSAY, J. and SILVERMAN,B. (2002). *A*pplied functional data analysis: Methods and Case Studies. Springer-Verlag, New York.

RAMSAY, J. and SILVERMAN,B. (2005). *F*unctional Data Analysis. Springer Science and Business Media, New York.

YAO, W. and LI, L. (2013). A new regression model: modal linear regression. *Scandinavian Journal of Statistics*, **41**, 656-671 .

Chapter 7

Conclusion and perspectives

In this thesis, we have seen the local linear estimation method for two different functions classes, the conditional distribution and the first derivative of regression function. When the explanatory variable is functional with a scalar response and in the iid setup, the uniform almost complete with convergence rates of two previous estimator functions is established where the convergence rate expressions of the conditional distribution function have the same as the conditional density function while the first derivative of regression function have the same rate expressions as the regression operator.

From several tools in nonparametric statistics, the conditional mode and the mode regression are studied in this thesis. The estimators of two latter tools are based on the bihaviour of conditional distribution function and regression function (their derivatives) respectively. Unfortunately, we think that much less is known about local linear mode regression for functional data, so this thesis is devoted to introduce some results which may contribute to inderstand and clarify the local linear mode regression estimator. However, these last results are not enough to give the really asymptotic behaviour framework of this novel therefore, we can propose the following perspectives :

- Studying the asymptotic normality of our estimtor in order to build confidence intervals.

- The mean squared error term of our estimator may contribute in the bandwidths choice (it may contribute to get an optimal bandwidth)
- When some experimental studies data are difficult to collect, we are looking forward to establish the almost complete convergence, the mean squared error and the asymptotic normality of our estimtor with responses missing at random (MAR).
- We can also generalize our results to the spatial framework.
- It is worth to study the asymptotic beheviour of our estimator with functional response.

Bibliography

ANEIROS-PÉREZ, G.,CAO, R. and VILAR-FERNÁNDEZ, J.M. (2011). Functional methods for time series prediction : A nonparametric approach. *J of Forecasting, **30**, 377-392.*

ATTAOUI, S.*,* LAKSACI, A. *and* OULD-SAÏD, E*. (2011). A note on the conditional density estimate in the single functional index model.* Statist. Probab. Lett, **81**(1), 45-53.

ATTOUCH, M., LAKSACI, A. and RAFAA, F.(2018). On the local linear estimate for functional regression : Uniform in bandwidth consistency. *Communications in Statistics-Theory and Methods, **48**, 1-18.*

AURZADA, F. *and* SIMON, T. *(2007). small ball probabilities for stable convolutions.* ESAIM Probab. Stat, **11**, 327-343 (electronic).

BAÍLLO, A. and GRANÉ, A. (2009). Functional Local Linear Regression with Functional Predictor and Scalar Response. *Journal of Multivariate Analysis, **100**, 102-111.*

BARRIENTOS-MARIN, J., FERRATY, F. *and* VIEU, P *.(2010). Locally Modelled Regression and Functional Data.*J. of Nonparametric Statistics, **22**, 617-632.

BENHENNI, K., FERRATY, F., RACHDI, M. AND VIEU, P. (2007). Local smoothing regression with functional data. *Computational Statistics. **22(3)**, 353- 369.*

BOJ, E. DELICADO, P. *and* FORTIANA, J. *(2010). Distance-based local linear regression for functional predictors.* Computational Statistics and Data Analysis, **54**, 429-437.

BOUANANI, O., LAKSACI, A., RACHDI, M. and RAHMANI, S. (2019). Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics. *Behaviormetrika, **46**, 199-233.*

BOULARAN, J., FERRÉ, L. AND VIEU, P*. (1995). Location of particular points in nonparametric regression analysis.* Austral. J. Statist. ***37**, 161-168.*

BORGGAARD, C., THODBERG, H.H.*(1992). Optimal Minimal Neural Interpretation of Spectra. Analytical Chemistry, **64**, 545-551.*

BOSQ., D *.(2000). Linear Processes in Function Spaces: Theory and Applications. Springer-Verlag.*

BURBA, F.;FERRATY, F. *and* VIEU, P. *(2009). k-Nearest Neighbour method in functional nonparametric regression, Journal of Nonparametric Statistics 21(4), 453-469 .*

CHEN, Y., MA, X. *and* ZHOU, J. *(2017). Variable selection for mode regression.* Journal of Applied Statistics, **45**, 1-8 .

COLLOMB, G.(1984). Propriétés de convergence presque compléte du prédicteur à noyau (in french).*Z. Wahrscheinlichkeitstheorie verw. Gebiete, **66**, 441-460.*

CRAMBES, C., KNEIP, A. *and* SARDA, P. *(2007). Smoothing splines estimators for functional linear regression,* Ann. Statist, **37**(1), 35-72.

CUEVAS, A. and FRAIMAN, R. (2004). On the bootstrap methodology for functional data, *COMPSTAT 2004-Proceedings in Computational Statistics, 127-135.*

DABO-NIANG, S. *and* RHOMARI, N. *(2003). Kernel regression estimation when the regressor takes values in metric space.* C. R. Acad. Sci. Paris*, **336**, 75-80.*

DAUXOIS, J., POUSSE, , A. *and* ROMAIN, Y. *(1982). Asymptotic theory for the principal component analysis of a vector random function : some applications to statistical inference,* J. Multivariate Anal, **12**(1), 136-154.

DELSOL, L. (2007a). CLT and $\mathbb{L}^q$ errors in nonparametric functional regression, *C. R. Math. Acad. Sci,* ***345*** *(7), 411-414.*

DELSOL, L. *(2007b). Régression non-paramétrique fonctionnelle : Expressions asymptotiques des moments,* Annales de l'I.S.U.P, **LI**(3), 43-67.

DELSOL, L. (2008). Régression sur variable fonctionnelle : Estimation, Tests de structure et Applications. PhD thesis from the Toulouse III - Paul Sabatier University (France).

DEMONGEOT, J., LAKSACI, A., MADANI, F. and RACHDI, M. (2011). *A fast functional locally modeled conditional density and mode for functional time-series.* Recent Advances in Functional Data Analysis and Related Topics, Contributions to Statistics, Physica-Verlag/Springer, 2011, Pages 85-90, DOI : $10.1007/978-3-7908-2736-113$.

DEMONGEOT, J., LAKSACI, A., MADANI, F. and RACHDI, M. (2013). Functional data : local linear estimation of the conditional density and its application.*Statistics*, **47**, 26-44.

DEMONGEOT, J., LAKSACI, A. RACHDI, M. and RAHMANI, S. (2014). On the local Modalization of the conditional distribution for functional data. *Sankhya A*, **76** (2), 328-355.

DEVILLE, JC. (1974). Méthodes statistiques et numériques de l'analyse harmonique. *Ann INSEE*, **15**, 3-101.

ERDÖS, P. (1949). On a theorem of Hsu and Robbins. *Ann. Math. Statist,* ***20*** *(2), 286-291.*

EZZAHRIOUI, M. *and* OULD-SAÏD,E. *(2008). Asymptotic normality of a nonpara-*

metric estimator of the conditional mode function for functional data. J Nonparametric Stat, **20**, 3-18.

Ezzahrioui, M. and Ould Saïd, E. (2010).Some asymptotic results of a nonparametric conditional mode estimator for functional time series data. *Neerlandica, **64**, 171-201.*

Fan, J. *and* Gijbels, I. *(1992). Variable Bandwidth and Local Linear Regression Smoothers.* Annals of Statistics, **20**, 2008-2036.

Fan, j. Gijbels, i. (1996). *Local Polynomial Modelling and its Applications.* Chapman & Hall, London.

Fan, J. and Yao, Q. (2003). *Nonlinear Time Series : Nonparametric and Parametric Methods. Springer-Verlag, New York.*

Ferraty, F. *and* Vieu, P. *(2000). Dimension fracale et estimation de la régression dans des espaces vectoriels semi-normés,* Compte Rendus de l'Académie des Sciences Paris, **330**, 403-406.

Ferraty, F. and Vieu, P. (2004). Nonparametric models for functional data, with application in regression, time series prediction and curve discrimination. *J. Nonparametric. Stat. **16** (1-2) 111-125.*

Ferraty, F.,Laksaci, A. *and*Vieu, P. *(2006). Estimating some characteristics of the conditional distribution in nonparametric functional models.* Statistical Inference for Stochastic Processes*, **9**, 47-76.*

Ferraty, F. *and* Vieu, P. *(2006). Nonparametric functional data analysis : Theory and Practice. Springer Series in Statistics, New York : Springer-Verlag.*

Ferraty, F., Laksaci, A., Tadj, A., *and* Vieu, P.*(2010). Rate of uniform consistency for nonparametric estimates with functional variables.* Journal of Statistical Planning and Inference*, **140**, 335-352.*

GAO, F. *and* LI, W.V. *(2007). Small ball probabilities for the Slepian Gaussian fields.* Transactions of the American Mathematical Society, **359** (3), 1339-1350. (electronic)

GYÖRFI, L., HÄRDLE, W., SARDA, P. and VIEU,P. (1989). Nonparametric Curve Estimation from Time Series.*Springer-Verlag New York.*

HALL, P. *and* HOSSEINI-NASAB, M. *(2006). On properties of functional principal components.* Journal of the Royal Statistical Society Series B, **68**, 109-126.

HASTIE, T. and MALLOWS, C. (1993). Discussion of "A statistical view of some chemometrics regression tools."by Frank, I.E. and Friedman, *J.H. Technometrics, **35**, 140-143.*

HEBCHI, C. *(2020). Uniform almost complete convergence of local linear mode regression,* IJSE, **21**(1), 54-62.

HONDA, T. (2000). Nonparametric estimation of a conditional quantile for α-mixing processes. *Annals of the Institute of Statistical Mathematics*, **52** (3), 459-470.

HSU, P.and ROBBINS, H. (1947). Complete convergence and the law of large numbers. *Proc. Nat. Acad. Sci. USA, **33**, 25-31.*

KARA, L-Z., LAKSACI, A., RACHDI, M. *and* VIEU, P. *(2017). Uniform in bandwidth consistency for various kernel estimators involving functional data.* Journal of Nonparametric Statistics, **29**, 85-107.

KEMP, G. and SANTOS SILVA, J. (2012). Regression towards the mode.*Journal of Econometrics, **170**, 92-101.*

KOENKER, R. *(2000). Galton, Edgeworth, Frisch, and prospects for quantile regression in econometrics.* Journal of Econometrics, **95**, 347-374.

KOENKER, R. (2005). A Frisch-Newton Algorithm for Sparse Quantile Regression.

Acta Mathematicae Applicatae Sinica, English Series, ***21****, 225-236.*

Koul, H. L. *and* Mukherjee. K. *(1994). Regression quantiles and related processes under long range dependent errors. Journal of Multivariate Analysis,* ***51****(2), 318-337.*

Laksaci, A. *(2007). Erreur quadratique de l'estimateur à noyau de la densité conditionnelle à variable explicative fonctionnelle.* C. R. Acad. Sci. Paris,***345****, 171-175.*

Laksaci, A., Lemdani, M. *and* Ould-Saïd, E. *(2008). A generalized L^1-approach for a kernel estimator of conditional quantile with functional regressors : consistency and asymptotic normality.*Stat. Probab. Lett, **79**, 1065-1073.

Laksaci, A., Madani, F. and Rachdi, M. (2012). Kernel conditional density estimation when the regressor is valued in a semi metric space. *Communications in Statistics-Theory and Methods,* ***42****(19), 3544-3570.*

Lee, M.J. *(1989). Mode regression.*J. Econometrics, **42**, .

Lee, M. J. (1993). Quadratic mode regression.*J. Econometrics,* ***57****, 1-19.*

Li, W.V. *and* Shao, Q.M. *(2001). Gaussian processes : inequalities, small ball probabilities and applications, In :C.R. Rao and D. Shanbhag (eds.) Stochastic processes, Theory and Methods. Handbook of Statitics, 19, North-Holland, Amsterdam.*

Ling, N. *and* Xu, Q. *(2012). Asymptotic normality of conditional density estimation in the single index model for functional time series data.* Statistics and Probability Letters, **82**, 2235-2243.

Mack ,Y. P. and Müller H. G. (1989). Derivative Estimation in Nonparametric Regression with Random Predictor Variable.*Sankhyā,* ***51****(1), 59-72.*

Madani, F. *(2012). Aspects théoriques et pratiques dans l'estimation non paramétrique* *de la densité conditionnelle pour des données fonctionnelles. PhD thesis from the Grenoble University (France).*

MAS, A. *(2012). Lower bound in regression for functional data by representation of small ball probabilities.* Electronic Journal of Statistics., ***6**, 1745-1778.*

MASRY, E.*(2005). Nonparametric regression estimation for dependent functional data : asymptotic normality. Stochastic Process and their Applications, **115** (1), 155-177.*

MESSACI, F., NEMOUCHI, N., OUASSOU, I. *and* RACHDI, M. *(2015). Local polynomial modelling of the conditional quantile for functional data.*Statistical Methods and Applications, **24**(4), 597-622.

PARZEN, E. (1962). Estimation of a probability density function and mode. *The Annals of Mathematical Statistics, **33**, 1065-1076.*

PONS, O. *(2011). Functional Estimation for Density, Regression Models and Processes. WORLD SCIENTIFIC, Singapore.*

PORTNOY, S. *(1991). Correction: Asymptotic Behavior of M Estimators of p Regression Parameters when p^2/n is Large: II. Normal Approximation.* Ann. Statist, **19**, 2282.

POWELL, J. (1986). Symmetrically trimmed least squares estimation for Tobit models. *Econometrica, **54**, 1435-1460.*

PREDA, C. *(1999). Analyse factorielle d'un processus : problème d'approximation et de régression (in french), PhD Lille I, 1999.*

PREDA, C. *and* SAPORTA, G. *(2002). Régression PLS sur un processus stochastique. Revue de Statistique Appliquée, **50** (2), 27-45.*

PREDA, C. *and* SAPORTA, G. *(2004). PLS approach for clusterwise linear regression on functional data. Classification, clustering, and data mining applications.*

167-176, Stud. Classification Data Anal. Knowledge Organ., Springer, Berlin.

Preda, C. *and* Saporta, G. *(2005a). PLS regression on a stochastic process.Comput.Statist. Data Anal,* ***48*** *(1), 149-158.*

Preda, C. *and* Saporta, G.*(2005b). Clusterwise PLS regression on a stochastic process. Comput.Statist. Data Anal,* ***49*** *(1), 99-108.*

Preda, C*. (2007). Regression models for functional data by reproducing kernel Hilbert spaces methods.* J. Statist. Plann. Inference, **137**(3), 829-840.

Rachdi, M., Laksaci, A.,Demongeot, J., Abdali, A. and Madani, F. (2014). Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data. *Computational Statistics and Data Analysis*, **73**, 53-68.

Ramsay, J. O. (1982). When data are functions. *Psychometrika,* ***47****(4), 379-396.*

Ramsay, J. *and* Dalzell, C*. (1991). Some tools for functional data analysis.*J. R. Statist. Soc. B, **53**, 539-572.

Ramsay, J. O. and Silverman, B.W.(1997). Functional data analysis. Springer series in statistics. Springer, New York.

Ramsay, J. O. and Silverman, B.W. (2002). Applied functional data analysis: Methods and Case Studies. Springer Series in Statistics. Springer-Verlag, New York.

Ramsay, J. O. and Silverman, B.W. (2005). Functional Data Analysis. 2nd ed.Springer, New-York.

Rosenblatt, M. (1969). M. Conditional probability density and regression estimators.*in Multivariate Analysis II, Ed. P.R. Krishnaiah,,* ***25-31****, New York: Academic Press.*

ROUSSEEUW, P. J. *(1984). Least Median of Squares Regression.* Journal of the American Statistical Association, **79**, 871-880 .

SAPORTA, G. (1981). Méthodes exploratoires d'analyse de données temporelles. *Cahiers du Bureau universitaire de Recherche Opérationnelle Série Recherche,* ***37****(38), 7-194.*

STONE, C. J. *(1977). Consistent Nonparametric Regression.* The Annals of Statistics, **5**(4), 595-620.

STUTE, W. (1986). Conditional Empirical Processes. *Ann. Statist,* ***14****, 638-647.*

USPENSKY, J.V. *(1937).* Introduction to Mathematical Probability*. McGraw-Hill Book Company.*

VALDERRAMA MJ. *(2007). An overview to modelling functional data.* Computational Statistics*,* ***22****, 331-334.*

WAND, M. *and* JONES, C.*(1995).* Kernel Smoothing, Monographs on Statistics and Applied Probability (Vol. 60), London : Chapman & Hall.

XIONG, X., ZHOU, P., AILIAN, CH. (2018). Asymptotic normality of the local linear estimation of the conditional density for functional time series data, *Communications in Statistics - Theory and Methods* **47**, 3418-3440 .

YAO, F. and LEE, T.C.M. (2006). Penalised spline models for functional principal component analysis, *Journal of the Royal Statistical Society Series B,* ***68*** *(1), 3-25.*

YAO, W. *and* LI, L. *(2013). A new regression model: modal linear regression.* Scandinavian Journal of Statistics, **41**, 656-671 .

YU, K., LU, Z. and STANDER,J. (2003). Quantile regression: applications and current research areas. *Journal of the Royal Statistical Society Series D*, **52** (3), 331-350.

ZHANG, J.T. (2014). *Analysis of Variance for Functional Data.* New York: Chapman and Hall/CRC.

ZHOU, Z. and LIN, Z. (2015). Asymptotic normality of locally modelled regression estimator for functional data, *Journal of Nonparametric Statistics, **28**, 1-16.*

ZIEGLER, K. *(2002). On nonparametric kernel estimation of the mode of the regression function in the random design model.* Journal of Nonparametric Statistics, **14**(6), 749-774.

ZUE, V., SENEFF, S., GLASS, J. (1990). Speech database development at MIT: TIMIT and BEYOND. Speech Communication, **9** 351-356